```
=> fil reg; d que l3; fil capl; d que l13
FILE 'REGISTRY' ENTERED AT 13:07:06 ON 13 DEC 2004
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2004 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file
provided by InfoChem.

STRUCTURE FILE UPDATES:   12 DEC 2004  HIGHEST RN 796841-97-9
DICTIONARY FILE UPDATES:  12 DEC 2004  HIGHEST RN 796841-97-9

TSCA INFORMATION NOW CURRENT THROUGH MAY 21, 2004

  Please note that search-term pricing does apply when
  conducting SmartSELECT searches.

Crossover limits have been increased.  See HELP CROSSOVER for details.

Experimental and calculated property data are now available. For more
information enter HELP PROP at an arrow prompt in the file or refer
to the file summary sheet on the web at:
http://www.cas.org/ONLINE/DBSS/registryss.html


L2        25918 SEA FILE=REGISTRY ABB=ON  .{3}C.{2}C.{12}H.{3}[-H].{4}/SQSP
L3          138 SEA FILE=REGISTRY ABB=ON  L2 AND SQL<50


FILE 'CAPLUS' ENTERED AT 13:07:06 ON 13 DEC 2004
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)


Copyright of the articles to which records in this database refer is
held by the publishers listed in the PUBLISHER (PB) field (available
for records published or updated in Chemical Abstracts after December
26, 1996), unless otherwise indicated in the original publications.
The CA Lexicon is the copyrighted intellectual property of the
American Chemical Society and is provided to assist you in searching
databases on STN.  Any dissemination, distribution, copying, or storing
of this information, without the prior written consent of CAS, is
strictly prohibited.

FILE COVERS 1907 - 13 Dec 2004  VOL 141 ISS 25
FILE LAST UPDATED: 12 Dec 2004  (20041212/ED)

  This file contains CAS Registry Numbers for easy and accurate
  substance identification.

'OBI' IS DEFAULT SEARCH FIELD FOR 'CAPLUS' FILE


L2        25918 SEA FILE=REGISTRY ABB=ON  .{3}C.{2}C.{12}H.{3}[-H].{4}/SQSP
L3          138 SEA FILE=REGISTRY ABB=ON  L2 AND SQL<50
L8           83 SEA FILE=CAPLUS ABB=ON  L3
L9           20 SEA FILE=CAPLUS ABB=ON  L8 NOT PY>1999
L12          35 SEA FILE=CAPLUS ABB=ON  L8 NOT AY>1999
L13          35 SEA FILE=CAPLUS ABB=ON  L9 OR L12
```

Searched by Barb O'Bryen, STIC 2-2518

=> d his ful ed abs hitstr 1-35

L13 ANSWER 1 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2004:984174 CAPLUS
TITLE: Genome sequence of Haloarcula marismortui: a halophilic archaeon from the Dead Sea
AUTHOR(S): Baliga, Nitin S.; Bonneau, Richard; Facciotti, Marc T.; Pan, Min; Glusman, Gustavo; Deutsch, Eric W.; Shannon, Paul; Chiu, Yulun; Weng, Rueyhung Sting; Gan, Rueichi Richie; Hung, Pingliang; Date, Shailesh V.; Marcotte, Edward; Hood, Leroy; Ng, Wailap Victor
CORPORATE SOURCE: Institute for Systems Biology, Seattle, WA, 98103, USA
SOURCE: Genome Research (2004), 14(11), 2221-2234
CODEN: GEREFS; ISSN: 1088-9051
PUBLISHER: Cold Spring Harbor Laboratory Press
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 17 Nov 2004
AB The complete sequence of the 4,274,642-bp genome of Haloarcula marismortui, a halophilic archaeal isolate from the Dead Sea, is reported. The genome is organized into 9 circular replicons of varying G+C compns. ranging from 54 to 62%. Comparison of the genome architectures of Halobacterium sp. NRC-1 and H. marismortui suggests a common ancestor for the two organisms and a genome of significantly reduced size in the former. Both of these halophilic archaea use the same strategy of high surface neg. charge of folded proteins as means to circumvent the salting-out phenomenon in a hypersaline cytoplasm. A multitiered annotation approach, including primary sequence similarities, protein family signatures, structure prediction, and a protein function association network, has assigned putative functions for at least 58% of the 4242 predicted proteins, a far larger number than is usually achieved in most newly sequenced microorganisms. Among these assigned functions were genes encoding 6 opsins, 19 MCP and/or HAMP domain signal transducers, and an unusually large number of environmental response regulators (nearly 5-fold more than those encoded in Halobacterium sp. NRC-1), suggesting H. marismortui is significantly more physiol. capable of exploiting diverse environments. In comparing the physiologies of the two halophilic archaea, in addition to the expected extensive similarity, several differences were discovered in their metabolic strategies and physiol. responses such as distinct pathways for arginine breakdown in each halophile. Finally, as expected from the larger genome, H. marismortui encodes many more functions and seems to have fewer nutritional requirements for survival than does Halobacterium sp. NRC-1. The two chromosome and 7 plasmid sequences are deposited in GenBank/EMBL/DDBJ under accession nos. AY596290-AY596296, and in the NCBI RefSeq database under accession nos. NC_006389-NC_006397.
IT 775160-56-0, GenBank AAV44349 775175-04-7, GenBank AAV45792 775190-06-2, GenBank AAV47294
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; complete genome sequence of Haloarcula marismortui, a halophilic archaeon from the Dead Sea)
REFERENCE COUNT: 64 THERE ARE 64 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

Use Registry # to match citation to sequence. Sequences are printed beginning on pg 27

L13 ANSWER 2 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2004:979095 CAPLUS
TITLE: The status, quality, and expansion of the NIH full-length cDNA project: The mammalian gene

collection (MGC)
AUTHOR(S): Gerhard, Daniela S.; Wagner, Lukas; Feingold, Elise A.; Shenmen, Carolyn M.; Grouse, Lynette H.; Schuler, Greg; Klein, Steven L.; Old, Susan; Rasooly, Rebekah; Good, Peter; Guyer, Mark; Peck, Allicon M.; Derge, Jeffery G.; Lipman, David; Collins, Francis S.
CORPORATE SOURCE: The MGC Project Team, NIH, USA
SOURCE: Genome Research (2004), 14(10b), 2121-2127
CODEN: GEREFS; ISSN: 1088-9051
PUBLISHER: Cold Spring Harbor Laboratory Press
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 17 Nov 2004
AB The National Institutes of Health's Mammalian Gene Collection (MGC) project was designed to generate and sequence a publicly accessible cDNA resource containing a complete open reading frame (ORF) for every human and mouse gene. The project initially used a random strategy to select clones from a large number of cDNA libraries from diverse tissues. Candidate clones were chosen based on 5'-EST sequences, and then fully sequenced to high accuracy and analyzed by algorithms developed for this project. Currently, more than 11,000 human and 10,000 mouse genes are represented in MGC by at least one clone with a full ORF. The random selection approach is now reaching a saturation point, and a transition to protocols targeted at the missing transcripts is now required to complete the mouse and human collections. Comparison of the sequence of the MGC clones to reference genome sequences reveals that most cDNA clones are of very high sequence quality, although it is likely that some cDNAs may carry missense variants as a consequence of exptl. artifact, such as PCR, cloning, or reverse transcriptase errors. Recently, a rat cDNA component was added to the project, and ongoing frog (Xenopus) and zebrafish (Danio) cDNA projects were expanded to take advantage of the high-throughput MGC pipeline. The sequence data for the full-length clones from this study have been submitted to GenBank/EMBL/DDBJ under accession nos. BC000001-BC077073. [This abstr record is one of 39 records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].
IT 623621-65-8, GenBank AAH46856
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; status, quality, and expansion of the NIH full-length cDNA project and mammalian gene collection (MGC))

L13 ANSWER 3 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2004:241810 CAPLUS
DOCUMENT NUMBER: 140:248280
TITLE: EST and contig sequences of Drosophila melanogaster and their uses in microarrays, retrieval of full-length cDNAs and proteomic analysis, and for identification of pesticide targets
INVENTOR(S): Homburger, Sheila Akiko; Ebens, Allen James, Jr.; Erickson, Catherine Sue; Francis-Lang, Helen Louise; Margolis, Jonathan Scott; Reddy, Bindu Priya; Ruddy, David Andrew; Buchman, Andrew Roy
PATENT ASSIGNEE(S): Exelixis, Inc., USA
SOURCE: U.S., 262 pp.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 19
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6703491	B1	20040309	US 1999-270767	19990317
US 6703491	B1	20040309	US 1999-270767	19990317

PRIORITY APPLN. INFO.: US 1999-270767 A 19990317

ED Entered STN: 24 Mar 2004

AB The present invention relates to Drosophila genes and methods for their use. A library of 31,629 expressed sequence tags and contig sequences are provided from tissues of mixed-stage embryos (0-20 h), imaginal disks, and adult heads of Drosophila melanogaster. Drosophila ESTs and sequence contigs derived from ESTs are useful as tools for retrieval of full-length protein coding sequences, for proteomic anal., for use in microarrays and gene expression anal., and for identification of pesticide targets. Thus, the invention provides nucleotide sequences of Drosophila genes, amino acid sequences of the encoded proteins, and derivs. (e.g., fragments) and analogs thereof. Special emphasis is given to DNA sequences encoding G protein-coupled receptors and chitin synthetase. The invention further relates to fragments (and derivs. and analogs thereof) of proteins which comprise one or more domains of a Drosophila protein. Antibodies to Drosophila proteins, and derivs. and analogs thereof, are also provided. Also provided herein are vectors and host cells comprising such nucleic acids. Methods of production of a Drosophila protein (e.g., by recombination means), and derivs. and analogs thereof, are provided. Chimeric polypeptide mols. comprising polypeptides of the invention fused to heterologous polypeptide sequences are provided. Methods to identify the biol. function of a Drosophila gene are provided, including various methods for the functional modification (e.g., overexpression, underexpression, mutation, knock-out) of one gene, or of two or more genes simultaneously. [This abstract record is one of sixteen records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].

IT 669850-32-2
RL: AGR (Agricultural use); BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)
(amino acid sequence; EST and contig sequences of Drosophila melanogaster and their uses in microarrays, retrieval of full-length cDNAs and proteomic anal., and for identification of pesticide targets)

L13 ANSWER 4 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2004:241804 CAPLUS
DOCUMENT NUMBER: 140:248276
TITLE: EST and contig sequences of Drosophila melanogaster and their uses in microarrays, retrieval of full-length cDNAs and proteomic analysis, and for identification of pesticide targets
INVENTOR(S): Homburger, Sheila Akiko; Ebens, Allen James, Jr.; Erickson, Catherine Sue; Francis-Lang, Helen Louise; Margolis, Jonathan Scott; Reddy, Bindu Priya; Ruddy, David Andrew; Buchman, Andrew Roy
PATENT ASSIGNEE(S): Exelixis, Inc., USA
SOURCE: U.S., 262 pp.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 19
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6703491	B1	20040309	US 1999-270767	19990317

US 6703491 B1 20040309 US 1999-270767 19990317
PRIORITY APPLN. INFO.: US 1999-270767 A 19990317
ED Entered STN: 24 Mar 2004
AB The present invention relates to Drosophila genes and methods for their use. A library of 31,629 expressed sequence tags and contig sequences are provided from tissues of mixed-stage embryos (0-20 h), imaginal disks, and adult heads of Drosophila melanogaster. Drosophila ESTs and sequence contigs derived from ESTs are useful as tools for retrieval of full-length protein coding sequences, for proteomic anal., for use in microarrays and gene expression anal., and for identification of pesticide targets. Thus, the invention provides nucleotide sequences of Drosophila genes, amino acid sequences of the encoded proteins, and derivs. (e.g., fragments) and analogs thereof. Special emphasis is given to DNA sequences encoding G protein-coupled receptors and chitin synthetase. The invention further relates to fragments (and derivs. and analogs thereof) of proteins which comprise one or more domains of a Drosophila protein. Antibodies to Drosophila proteins, and derivs. and analogs thereof, are also provided. Also provided herein are vectors and host cells comprising such nucleic acids. Methods of production of a Drosophila protein (e.g., by recombination means), and derivs. and analogs thereof, are provided. Chimeric polypeptide mols. comprising polypeptides of the invention fused to heterologous polypeptide sequences are provided. Methods to identify the biol. function of a Drosophila gene are provided, including various methods for the functional modification (e.g., overexpression, underexpression, mutation, knock-out) of one gene, or of two or more genes simultaneously. [This abstract record is one of sixteen records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].
IT 669250-26-4, 669255-25-8
RL: AGR (Agricultural use); BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)
(amino acid sequence; EST and contig sequences of Drosophila melanogaster and their uses in microarrays, retrieval of full-length cDNAs and proteomic anal., and for identification of pesticide targets)

L13 ANSWER 5 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2004:240442 CAPLUS
DOCUMENT NUMBER: 140:248267
TITLE: EST and contig sequences of Drosophila melanogaster and their uses in microarrays, retrieval of full-length cDNAs and proteomic analysis, and for identification of pesticide targets
INVENTOR(S): Homburger, Sheila Akiko; Ebens, Allen James, Jr.; Erickson, Catherine Sue; Francis-lang, Helen Louise; Margolis, Jonathan Scott; Reddy, Bindu Priya; Ruddy, David Andrew; Buchman, Andrew Roy
PATENT ASSIGNEE(S): Exelixis, Inc., USA
SOURCE: U.S., 262 pp.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 19
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6703491	B1	20040309	US 1999-270767	19990317
US 6703491	B1	20040309	US 1999-270767	19990317

PRIORITY APPLN. INFO.: US 1999-270767 A 19990317
ED Entered STN: 24 Mar 2004

AB The present invention relates to Drosophila genes and methods for their use. A library of 31,629 expressed sequence tags and contig sequences are provided from tissues of mixed-stage embryos (0-20 h), imaginal disks, and adult heads of Drosophila melanogaster. Drosophila ESTs and sequence contigs derived from ESTs are useful as tools for retrieval of full-length protein coding sequences, for proteomic anal., for use in microarrays and gene expression anal., and for identification of pesticide targets. Thus, the invention provides nucleotide sequences of Drosophila genes, amino acid sequences of the encoded proteins, and derivs. (e.g., fragments) and analogs thereof. Special emphasis is given to DNA sequences encoding G protein-coupled receptors and chitin synthetase. The invention further relates to fragments (and derivs. and analogs thereof) of proteins which comprise one or more domains of a Drosophila protein. Antibodies to Drosophila proteins, and derivs. and analogs thereof, are also provided. Also provided herein are vectors and host cells comprising such nucleic acids. Methods of production of a Drosophila protein (e.g., by recombination means), and derivs. and analogs thereof, are provided. Chimeric polypeptide mols. comprising polypeptides of the invention fused to heterologous polypeptide sequences are provided. Methods to identify the biol. function of a Drosophila gene are provided, including various methods for the functional modification (e.g., overexpression, underexpression, mutation, knock-out) of one gene, or of two or more genes simultaneously. [This abstract record is one of sixteen records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].

IT [redacted]

RL: AGR (Agricultural use); BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)

(amino acid sequence; EST and contig sequences of Drosophila melanogaster and their uses in microarrays, retrieval of full-length cDNAs and proteomic anal., and for identification of pesticide targets)

L13 ANSWER 6 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2004:209218 CAPLUS
DOCUMENT NUMBER: 140:230622
TITLE: Soybean nucleic acids and encoded proteins associated with transcription in plants and their uses for plant improvement
INVENTOR(S): La Rosa, Thomas J.; Zhou, Yihua; Kovalic, David K.; Cao, Yongwei
PATENT ASSIGNEE(S): USA
SOURCE: U.S. Pat. Appl. Publ., 15 pp., Cont.-in-part of U.S. Ser. No. 985,678, abandoned.
CODEN: USXXCO
DOCUMENT TYPE: Patent
LANGUAGE: English
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2004031072 A1		20040212	US 2003-XR424599	20030428
PRIORITY APPLN. INFO.:			US 1999-304517	19990506
			US 2001-985678	20011105
			US 2003-424599	20030428

ED Entered STN: 16 Mar 2004

AB This invention provides 142,842 polynucleotide sequences isolated from a cDNA library generated from Glycine maximum The open reading frame in each polynucleotide sequence is identified by a combination of predictive and homol.-based methods. Functions of polypeptides encoded by the polynucleotides sequences are determined using a hierarchical classification

tool, termed FunCAT, for Functional Categories Annotation Tool. Sequences useful for producing transgenic plants having improved biol. properties are identified from their FunCAT annotations. [This abstract record is one of 72 records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].

IT 666902-20-1
RL: BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)
(amino acid sequence; soybean nucleic acids and encoded proteins associated with transcription in plants and their uses for plant improvement)

L13 ANSWER 7 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2004:209164 CAPLUS
DOCUMENT NUMBER: 140:212085
TITLE: Soybean nucleic acids and encoded proteins associated with transcription in plants and their uses for plant improvement
INVENTOR(S): La Rosa, Thomas J.; Zhou, Yihua; Kovalic, David K.; Cao, Yongwei
PATENT ASSIGNEE(S): USA
SOURCE: U.S. Pat. Appl. Publ., 15 pp., Cont.-in-part of U.S. Ser No. 985,678, abandoned.
CODEN: USXXCO
DOCUMENT TYPE: Patent
LANGUAGE: English
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2004031072 A1		20040212	US 2003-XI424599	20030428
PRIORITY APPLN. INFO.:			US 1999-304517	19990506
			US 2001-985678	20011105
			US 2003-424599	20030428

ED Entered STN: 16 Mar 2004
AB This invention provides 142,842 polynucleotide sequences isolated from a cDNA library generated from Glycine maximum. The open reading frame in each polynucleotide sequence is identified by a combination of predictive and homol.-based methods. Functions of polypeptides encoded by the polynucleotides sequences are determined using a hierarchical classification tool, termed FunCAT, for Functional Categories Annotation Tool. Sequences useful for producing transgenic plants having improved biol. properties are identified from their FunCAT annotations. [This abstract record is one of 72 records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].

IT 665228-20-6
RL: BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses)
(amino acid sequence; soybean nucleic acids and encoded proteins associated with transcription in plants and their uses for plant improvement)

L13 ANSWER 8 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2004:116410 CAPLUS
DOCUMENT NUMBER: 140:247982
TITLE: Comparative genomic analysis of hyperthermophilic archaeal Fuselloviridae viruses
AUTHOR(S): Wiedenheft, Blake; Stedman, Kenneth; Roberto, Francisco; Willits, Deborah; Gleske, Anne-Kathrin;

Zoeller, Luisa; Snyder, Jamie; Douglas, Trevor; Young, Mark
CORPORATE SOURCE: Thermal Biology Institute, Montana State University, Bozeman, MT, 59717, USA
SOURCE: Journal of Virology (2004), 78(4), 1954-1961
CODEN: JOVIAM; ISSN: 0022-538X
PUBLISHER: American Society for Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 13 Feb 2004
AB The complete genome sequences of two Sulfolobus spindle-shaped viruses (SSVs) from acidic hot springs in Kamchatka (Russia) and Yellowstone National Park (United States) have been determined These nonlytic temperate viruses were isolated from hyperthermophilic Sulfolobus hosts, and both viruses share the spindle-shaped morphol. characteristic of the Fuselloviridae family. These two genomes, in combination with the previously determined SSV1 genome from Japan and the SSV2 genome from Iceland, have allowed us to carry out a phylogenetic comparison of these geog. distributed hyperthermal viruses. Each virus contains a circular double-stranded DNA genome of .apprx.15 kbp with approx. 34 open reading frames (ORFs). These Fusellovirus ORFs show little or no similarity to genes in the public databases. In contrast, 18 ORFs are common to all four isolates and may represent the minimal gene set defining this viral group. In general, ORFs on one half of the genome are colinear and highly conserved, while ORFs on the other half are not. One shared ORF among all four genomes is an integrase of the tyrosine recombinase family. All four viral genomes integrate into their host tRNA genes. The specific tRNA gene used for integration varies, and one genome integrates into multiple loci. Several unique ORFs are found in the genome of each isolate.
IT 606872-49-5
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; comparative genomic anal. of hyperthermophilic archaeal Fuselloviridae viruses)
REFERENCE COUNT: 44 THERE ARE 44 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L13 ANSWER 9 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2003:428587 CAPLUS
DOCUMENT NUMBER: 140:23775
TITLE: Relationships between fuselloviruses infecting the extremely thermophilic archaeon Sulfolobus: SSV1 and SSV2
AUTHOR(S): Stedman, Kenneth M.; She, Qunxin; Phan, Hien; Arnold, Hans Peter; Holz, Ingelore; Garrett, Roger A.; Zillig, Wolfram
CORPORATE SOURCE: Max-Planck-Institut fur Biochemie, Martinsried, 82152, Germany
SOURCE: Research in Microbiology (2003), 154(4), 295-302
CODEN: RMCREW; ISSN: 0923-2508
PUBLISHER: Editions Scientifiques et Medicales Elsevier
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 04 Jun 2003
AB The fusellovirus SSV2 from an Icelandic Sulfolobus strain was isolated, characterized and its complete genomic sequence determined SSV2 is very similar in morphol., replication, genome size and number of open reading frames (ORFs) to the type virus of the family, SSV1 from Japan, except in its high level of uninduced virus production The nucleotide sequences are, however, only 55% identical to each other, much less than related bacteriophage, related animal viruses and the rudiviruses of Sulfolobus,

SIRV1 and SIRV2. Nevertheless the genome architecture is very similar between the two viruses, indicating that despite this genomic dissimilarity the virus genomes are mostly homologous. Unlike SSV1, the sequence of SSV2 indicates integration into a glycyl tRNA gene and is completely missing a DNA packaging gene. There is a unique, perfectly tandemly directly repeated sequence of 62 nucleotides in SSV2 that has no similarity to known sequences or structures. By comparison to the SSV2 genome, an integrated partial fusellovirus genome was found in the Sulfolobus solfataricus P2 genome further confirming the dynamism of the Sulfolobus genome. Clustering of cysteine codon containing ORFs both in SSV1 and SSV2 indicates that these Fuselloviridae arose from a genome fusion event.

IT 583804-75-5, GenBank AAQ73267
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; relationships between fusselloviruses infecting the extremely thermophilic archaeon Sulfolobus, SSV1 and SSV2)

REFERENCE COUNT: 41 THERE ARE 41 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L13 ANSWER 10 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2002:357892 CAPLUS
DOCUMENT NUMBER: 137:105376
TITLE: Novel Strategy for the Design of a New Zinc Finger: Creation of a Zinc Finger for the AT-Rich Sequence by α-Helix Substitution
AUTHOR(S): Nagaoka, Makoto; Doi, Yoshihide; Kuwahara, Jun; Sugiura, Yukio
CORPORATE SOURCE: Institute for Chemical Research, Kyoto University, Uji, Kyoto, 611-0011, Japan
SOURCE: Journal of the American Chemical Society (2002), 124(23), 6526-6527
CODEN: JACSAT; ISSN: 0002-7863
PUBLISHER: American Chemical Society
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 14 May 2002
AB In this communication, a novel strategy for the design of a zinc finger peptide on the basis of α-helix substitution has been demonstrated. Sp1HM is a helix-substituted mutant for the wild-type Sp1(zf123) and its α-helix of each finger is replaced by that of fingers 4-6 of CF2-II. The CD spectrum of Sp1HM suggests that Sp1HM has an ordered secondary structure similar to that of Sp1(zf123). From the analyses of the DNA binding affinity and specificity by gel mobility shift assay, it is clearly indicated that Sp1HM specifically binds to the AT-rich sequence (5'-GTA TAT ATA-3') with 3.2 nM dissociation consts. Moreover, the zinc finger peptides for the sequence alternating between the AT- and GC-rich subsites can also be created by the α-helix substitution. This strategy is evidently effective and is also more convenient than the phage display method. Consequently, our design method is widely applicable to creating zinc finger peptides with novel binding specificities.
IT 443102-51-0
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(creation of a zinc finger for AT-rich sequence by α-helix substitution)

REFERENCE COUNT: 33 THERE ARE 33 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L13 ANSWER 11 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2000:540554 CAPLUS

DOCUMENT NUMBER: 133:130477
TITLE: The genome sequence of the plant pathogen Xylella fastidiosa
AUTHOR(S): Simpson, A. J. G.; Reinach, F. C.; Arruda, P.; Abreu, F. A.; Acencio, M.; Alvarenga, R.; Alves, L. M. C.; Araya, J. E.; Bala, G. S.; Baptista, C. S.; Barros, M. H.; Bonaccorsi, E. D.; Bordin, S.; Bove, J. M.; Briones, M. R. S.; Bueno, M. R. P.; Camargo, A. A.; Camargo, L. E. A.; Carraro, D. M.; Carrer, H.; Colauto, N. B.; Colombo, C.; Costa, F. F.; Costa, M. C. R.; Costa-Neto, C. M.; Coutinho, L. L.; Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A. P.; Ferreira, A. J. S.; Ferreira, V. C. A.; Ferro, J. A.; Fraga, J. S.; Franca, S. C.; Franco, M. C.; Frohme, M.; Furtan, L. R.; Garnier, M.; Goldman, G. H.; Goldman, M. H. S.; Gomes, S. L.; Gruber, A.; Ho, P. L.; Hoheisel, J. D.; Junqueira, M. L.; Kemper, E. L.; Kitajima, J. P.; Kreiger, J. E.; Duramae, E. E.; Laigret, F.; Lambals, M. R.; Lette, L. C. C.; Lemos, E. G. M.; Lemos, M. V. F.; Lopes, S. A.; Lopes, C. R.; Machado, J. A.; Machado, M. A.; Madeira, A. M. B. N.; Madeira, H. M. F.; Marino, C. L.; Marques, M. V.; Martins, E. A. L.; Martins, E. M. F.; Matsukuma, A. Y.; Menck, C. F. M.; Miracca, E. C.; Miyaki, C. Y.; Monteiro-Vitorello, C. B.; Moon, D. H.; Nagai, M. A.; Nascimento, A. L. T. O.; Netto, L. E. S.; Nhanl, A., Jr.; Nobrega, F. G.; Nunes, L. R.; Oliveira, M. A.; de Oilveria, M. C.; de Oliveira, R. C.; Palmieri, D. A.; Paris, A.; Peixoto, B. R.; Pereira, G. A. G.; Perelra, H. A.; Pesquero, J. B.; Quaggio, R. B.; Roberto, P. G.; Rodrigues, V.; Rosa, A. J. de M.; de Rosa, V. E., Jr.; de Sa, R. G.; Santelli, R. V.; Sawasaki, H. E.; da Silva, A. C. R.; da Silva, A. M.; da Silva, F. R.; Silva, W. A., Jr.; da Silveira, J. F.; Silvestri, M. L. Z.; Siqueira, W. J.; de Souza, A. A.; de Souza, A. P.; Terenzi, M. F.; Truffi, D.; Tsai, S. M.; Tsuhako, M. H.; Vallada, H.; Van Sluys, M. A.; Verjovski-Almeida, S.; Vettore, A. L.; Zago, M. A.; Zatz, M.; Meidanis, J.; Setubal, J. C.
CORPORATE SOURCE: Instituo Ludwig de Pesquisa sobre o Cancer, Sao Paulo, 01509-010, Brazil
SOURCE: Nature (London) (2000), 406(6792), 151-157
CODEN: NATUAS; ISSN: 0028-0836
PUBLISHER: Nature Publishing Group
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 08 Aug 2000
AB Xylella fastidiosa is a fastidious, xylem-limited bacterium that causes a range of economically important plant diseases. The complete genome sequence of X. fastidiosa clone 9a5c, which causes citrus variegated chlorosis-a serious disease of orange trees, is reported. The genome comprises a 52.7% GC-rich 2,679,305-base-pair (bp) circular chromosome and two plasmids of 51,158 bp and 1,285 bp. Putative functions can be assigned to 47% of the 2904 predicted coding regions. Efficient metabolic functions are predicted, with sugars as the principal energy and carbon source, supporting existence in the nutrient-poor xylem sap. The mechanisms associated with pathogenicity and virulence involve toxins, antibiotics and ion sequestration systems, as well as bacterium-bacterium and bacterium-host interactions mediated by a range of proteins. Orthologs of some of these proteins have only been identified in animal

and human pathogens; their presence in X. fastidiosa indicates that the mol. basis for bacterial pathogenicity is both conserved and independent of host. At least 83 genes are bacteriophage-derived and include virulence-associated genes from other bacteria, providing direct evidence of phage-mediated horizontal gene transfer.

IT 285136-62-1
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence; genome sequence of the plant pathogen Xylella fastidiosa)

L13 ANSWER 12 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2000:384548 CAPLUS
DOCUMENT NUMBER: 133:39116
TITLE: Genes and polypeptides involved in insulin signaling pathways for glucose tolerance, obesity, and longevity and their uses as therapeutic and diagnostic tools
INVENTOR(S): Ruvkun, Gary; Ogg, Scott
PATENT ASSIGNEE(S): The General Hospital Corporation, USA
SOURCE: PCT Int. Appl., 402 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 2
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000033068	A1	20000608	WO 1999-US28529	19991202
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 2001029617	A1	20011011	US 1998-205658	19981203
EP 1163515	A1	20011219	EP 1999-960641	19991202
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				

PRIORITY APPLN. INFO.:		
US 1998-205658	A	19981203
WO 1998-US10080	W	19980515
WO 1999-US28529	W	19991202

ED Entered STN: 09 Jun 2000
AB Disclosed herein are novel genes and methods for the screening of therapeutics useful for treating impaired glucose tolerance conditions, as well as diagnostics and therapeutic compns. for identifying or treating such conditions. The Caenorhabditis elegans metabolic regulatory genes daf-2 and age-1 encode homologs of the mammalian insulin receptor/phosphoinositol 3-kinase signaling pathway proteins, resp. Also, the C. elegans PKB kinase and AKT kinase act downstream of these genes, as their mammalian homologs act downstream of insulin signaling. The C. elegans PTEN lipid phosphatase homolog, DAF-18, acts upstream of AKT in this signaling pathway. Further, the DAF-16 forkhead protein represents the major transcriptional output of this insulin signaling pathway. Addnl. evidence indicates that the DAF-16, DAF-3, DAF-8, and DAF-14 transcriptional outputs of converging signaling pathways regulate metabolism The congruence between the C. elegans and mammalian insulin signaling pathways strongly supports the contention that new genes identified in the

C. elegans pathway also act in mammalian insulin signaling. Exemplary sequences and functional characteristics of the C. elegans daf genes and their human homologs are provided.

IT 274261-67-5 274262-13-4
RL: PRP (Properties)
(unclaimed protein sequence; genes and polypeptides involved in insulin signaling pathways for glucose tolerance, obesity, and longevity and their uses as therapeutic and diagnostic tools)

REFERENCE COUNT: 6 THERE ARE 6 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L13 ANSWER 13 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 2000:241545 CAPLUS
DOCUMENT NUMBER: 132:275965
TITLE: Divalent cation-independent cleavage of nucleic acids by zinc finger peptides
INVENTOR(S): Lima, Walt F.; Crooke, Stanley T.; Manoharan, Muthiah
PATENT ASSIGNEE(S): Isis Pharmaceuticals, Inc., USA
SOURCE: PCT Int. Appl., 38 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000020622	A1	20000413	WO 1999-US23273	19991006
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 6365379	B1	20020402	US 1999-412499	19991005
EP 1121455	A1	20010808	EP 1999-954760	19991006
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				
JP 2002526118	T2	20020820	JP 2000-574714	19991006
JP 3516920	B2	20040405		

PRIORITY APPLN. INFO.:	US 1998-103309P	P	19981006
	WO 1999-US23273	W	19991006

ED Entered STN: 14 Apr 2000
AB Selective cleavage of single stranded nucleic acids can be effected by contacting the nucleic acid with a zinc finger peptide in dimeric form. Dimerization results from diminution or elimination of zinc from the peptide, such that easily controllable and highly selective cleavage may be realized. Specificity may be obtained by conjugating the peptide with an oligonucleotide that will guide the peptide to a target sequence. Mechanisms of conjugating the peptides to oligonucleotides are described. The cleavage reaction shows Michaelis-Menten kinetics and the broad pH curve typical of acid-base catalyzed hydrolysis of RNA.

IT 136444-42-3D, dimers
RL: CAT (Catalyst use); PRP (Properties); USES (Uses)
(divalent cation-independent cleavage of nucleic acids by zinc finger peptides)

REFERENCE COUNT: 3 THERE ARE 3 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L13 ANSWER 14 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1999:622318 CAPLUS
DOCUMENT NUMBER: 131:239486
TITLE: Cloning, sequence and expression of zinc-binding LIM domain-containing protein S2-6
INVENTOR(S): Lecka-Czernik, Beata
PATENT ASSIGNEE(S): The Board of Trustees of the University of Arkansas, USA
SOURCE: U.S., 31 pp.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5959081	A	19990928	US 1997-856444	19970514
PRIORITY APPLN. INFO.:			US 1997-856444	19970514

ED Entered STN: 30 Sep 1999
AB A substantially pure S2-6 protein (a) having a zinc-binding LIM domain and nucleic acids encoding the S2-6 are disclosed. The S2-6 protein mRNA is preferentially expressed in nonproliferating or growth inhibited human diploid fibroblasts and overexpressed in senescent human diploid fibroblasts or human diploid fibroblasts derived from a patient with Werner syndrome. The S2-6 mRNA expression is reduced or abolished in fetal human diploid fibroblasts, immortalized cells, cancerous cells and other highly proliferative cells. The S2-6 protein may be an inhibitor of DNA synthesis in nonproliferating cells. The S2-6 protein may play a role in regulation of cell growth and differentiation.
IT 224425-86-9
RL: PRP (Properties)
(Unclaimed; cloning, sequence and expression of zinc-binding LIM domain-containing protein S2-6)
REFERENCE COUNT: 1 THERE ARE 1 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L13 ANSWER 15 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1999:593300 CAPLUS
DOCUMENT NUMBER: 131:319593
TITLE: Highly efficient endonucleolytic cleavage of RNA by a Cys2His2 zinc-finger peptide
AUTHOR(S): Lima, Walt F.; Crooke, Stanley T.
CORPORATE SOURCE: Department of Molecular and Structural Biology, Isis Pharmaceuticals, Carlsbad, CA, 92008, USA
SOURCE: Proceedings of the National Academy of Sciences of the United States of America (1999), 96(18), 10010-10015
CODEN: PNASA6; ISSN: 0027-8424
PUBLISHER: National Academy of Sciences
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 21 Sep 1999
AB We have identified a 30-aa peptide that efficiently cleaves single-stranded RNA. The peptide sequence corresponds to a single zinc finger of the human male-associated ZFY protein; a transcription factor belonging to the Cys2His2 family of zinc-finger proteins. RNA cleavage was observed only in the absence of zinc. Coordination with zinc resulted in complete loss of RNase activity. The RNase active structure was determined to be a homodimeric form of the peptide. Dimerization of the peptide occurred through a single intermol. disulfide between two of the four

cystines. The observed hydrolytic activity was single-stranded RNA-specific. Single-stranded DNA, double-stranded RNA and DNA, and 2'-methoxy-modified sequences were not degraded by the peptide. The peptide specifically cleaved pyrimidines within single-stranded RNA and the dinucleotide sequence 5'-pyr-A-3' was preferred. The RNA cleavage products consisted of a 3' phosphate and 5' hydroxyl. The initial rates of cleavage (V0) observed for the finger peptide were comparable to rates observed for human RNases, and the catalytic rate (Kcat) was comparable to rates observed for the group II intron ribozymes. The pH profile exhibited by the peptide is characteristic of general acid-base catalytic mechanisms observed with other RNases. These observations raise interesting questions about the potential biol. roles of zinc-finger proteins.

IT 136444-42-3D, disulfide-linked dimer
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(highly efficient endonucleolytic cleavage of RNA by a Cys2His2 zinc-finger peptide)

REFERENCE COUNT: 27 THERE ARE 27 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L13 ANSWER 16 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1999:330062 CAPLUS
DOCUMENT NUMBER: 130:348209
TITLE: Human zinc finger-containing DNA-binding protein S1-3 involved in inhibition of cell growth and its cDNA
INVENTOR(S): Lecka-Czernik, Beata
PATENT ASSIGNEE(S): University of Arkansas, USA
SOURCE: U.S., 49 pp.
CODEN: USXXAM
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 5905146	A	19990518	US 1996-616857	19960315

PRIORITY APPLN. INFO.: US 1996-616857 19960315

ED Entered STN: 28 May 1999

AB S1-3 protein, a human DNA binding protein containing three zinc finger domains and nucleic acids encoding S1-3 are disclosed. S1-3 mRNA is overexpressed in senescent human diploid fibroblasts and human diploid fibroblasts derived from a patient with Werner Syndrome but is not expressed in fetal human diploid fibroblasts. Microinjection of S1-3 antisense of S1-3 partial sense RNA into non-proliferating human fibroblasts stimulated DNA synthesis. S1-3 protein bound specifically to DNA and its binding site consensus sequence (GATRRWWG; R=purine, W=A, T) is found in many origins of DNA replication and overlaps a number of defined DNA binding sites for major transcription factors (GATA-1, NF-E1, AP1, and E2A) that have established function in cell proliferation and differentiation. The cDNA for S1-3 was identified by constructing and screening a subtracted cDNA library derived from polyA RNA of prematurely senescent Werner syndrome (WS) human diploid fibroblasts (HDF). Many cDNA clones representing genes overexpressed in senescent and WS HDF were identified. CDNAs representing six known genes coding for acid sphingomyelinase, fibronectin, SPARC, nm23-metastasis suppressor protein, and translation factors eIF-2β and EF-1a were found. Among the clones with no homol. to known sequences was clone S1-3.

IT **224425-86-9**
RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)

(S1-3 zinc finger domain 2; human zinc finger-containing DNA-binding protein S1-3 involved in inhibition of cell growth and its cDNA)
REFERENCE COUNT: 24 THERE ARE 24 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L13 ANSWER 17 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1999:113804 CAPLUS
DOCUMENT NUMBER: 130:192760
TITLE: 5'-Expressed sequence tags for secreted proteins identified from human brain tissues
INVENTOR(S): Dumas Milne Edwards, Jean-Baptiste; Duclert, Aymeric; Lacroix, Bruno
PATENT ASSIGNEE(S): Genset, Fr.
SOURCE: PCT Int. Appl., 581 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9906552	A2	19990211	WO 1998-IB1236	19980731
WO 9906552	A3	19990422		
W: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, SD, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				
US 6222029	B1	20010424	US 1997-905223	19970801
AU 9885555	A1	19990222	AU 1998-85555	19980731
EP 1000150	A2	20000517	EP 1998-936594	19980731
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI				
JP 2001512015	T2	20010821	JP 2000-505293	19980731
PRIORITY APPLN. INFO.:			US 1997-905223 A	19970801
			WO 1998-IB1236 W	19980731

ED Entered STN: 19 Feb 1999
AB The sequences of the 5' ends of 233 expressed sequence tags (ESTs) derived from mRNAs encoding human secreted proteins expressed in the brain are disclosed. Chemical and enzymic methods of obtaining mRNAs with intact 5' ends are described: (1) a chemical modification method involving derivatization of the 5' ends of the mRNAs and selection of the derivatized mRNAs, and (2) treatment with alkaline phosphatase to remove the phosphate groups present on the 5' ends of uncapped incomplete mRNAs and treatment with a decapping enzyme such as T4 polynucleotide kinase or tobacco acid pyrophosphatase to remove the cap present on full-length mRNAs. The 5' ESTs are screened to identify those having an uninterrupted open reading frame longer than 45 nucleotides beginning with an ATG codon and extending to the end of the EST, and then search to identify potential signal motifs with a score of at least 3.5 in the Von Heijne signal peptide identification matrix. The 5' ESTs may be to obtain cDNAS, and genomic DNAs corresponding to the 5' ESTs. The 5' ESTs may also be used in diagnostic, forensic, gene therapy, and chromosome mapping procedures. Upstream regulatory sequences may also be obtained using the 5' ESTs. The 5' ESTs may also be used to design expression vectors and secretion vectors.
IT 220663-89-8P

RL: BPN (Biosynthetic preparation); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation); USES (Uses) (amino acid sequence; 5'-ends of expressed sequence tags for secreted proteins from human brain tissues)

L13 ANSWER 18 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1997:640824 CAPLUS
DOCUMENT NUMBER: 127:302663
TITLE: Peptidyl fluorescent chemosensor for divalent zinc
INVENTOR(S): Imperiali, Barbara; Walkup, Grant K.
PATENT ASSIGNEE(S): California Institute of Technology, USA
SOURCE: PCT Int. Appl., 35 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9735182	A1	19970925	WO 1997-US4672	19970321
W: JP				
RW: AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE				
US 5928955	A	19990727	US 1996-620151	19960322
EP 888533	A1	19990107	EP 1997-917582	19970321
R: DE, ES, FR, GB, IE				
PRIORITY APPLN. INFO.:			US 1996-620151 A	19960322
			WO 1997-US4672 W	19970321

OTHER SOURCE(S): MARPAT 127:302663
ED Entered STN: 09 Oct 1997
AB The present invention provides a selective fluorescent chemosensor, sensitive to nanomolar concns. of zinc (II) and selective for this ion over Na+, Mg2+, Co2+.
IT 197169-20-3
RL: RCT (Reactant); RACT (Reactant or reagent)
(divalent zinc determination by fluorescent chemosensor based on zinc binding peptide covalently attached to fluorescent reporting group)

L13 ANSWER 19 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1996:408726 CAPLUS
DOCUMENT NUMBER: 125:135153
TITLE: Molecular cloning and sequencing of cDNAs encoding insecticidal peptides from the primitive hunting spider, Plectreurys tristis (Simon)
AUTHOR(S): Leisy, Douglas J.; Mattson, Jeanine D.; Quistad, Gary B.; Kramer, Steven J.; Van Beek, Nikolai; Tsai, Leslie W.; Enderlin, Frances E.; Woodworth, Alison R.; Digan, Mary Ellen
CORPORATE SOURCE: Sandoz Agro Inc., Palo Alto, CA, 94304, USA
SOURCE: Insect Biochemistry and Molecular Biology (1996), 26(5), 411-417
CODEN: IBMBES; ISSN: 0965-1748
PUBLISHER: Elsevier
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 13 Jul 1996
AB Plectreurys tristis cephalothorax mRNA was isolated and amplified by PCR using degenerate primers corresponding to reverse translated mature Plt-VI toxin. An oligonucleotide corresponding to a portion of the amplified product was then used to screen a P. tristis cDNA library. The cDNAs from 10 pos. clones were sequenced. Eight of these cDNAs corresponded to

Plt-VI toxin, one to Plt-XI toxin, and one was very similar to Plt-VIII toxin, with the exception of a single amino acid substitution. Anal. of these cDNAs indicated that these toxins are initially synthesized as prepro-forms which undergo signal cleavage followed by addnl. processing at both their N- and C-termini to produce the mature products.

IT 151066-05-6 151066-09-0 179986-33-5
RL: BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study); OCCU (Occurrence)
(amino acid sequence; mol. cloning and sequencing of cDNAs encoding insecticidal peptides from the primitive hunting spider, Plectreurys tristis (Simon))

L13 ANSWER 20 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1994:428898 CAPLUS
DOCUMENT NUMBER: 121:28898
TITLE: Isolation and sequencing of insecticidal peptides from the primitive hunting spider, Plectreurys tristis (Simon)
AUTHOR(S): Quistad, Gary B.; Skinner, Wayne S.
CORPORATE SOURCE: Sandoz Agro, Inc., Palo Alto, CA, 94304, USA
SOURCE: Journal of Biological Chemistry (1994), 269(15), 11098-101
CODEN: JBCHA3; ISSN: 0021-9258
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 23 Jul 1994
AB About 50 peptide toxins were purified from venom of the primitive hunting spider, Plectreurys tristis. Bioassay by injection into larval Heliothis virescens (tobacco budworm) allowed selection of nine toxins for anal. of amino acid sequences. Total sequences were determined for six of the more insecticidal peptides (46-49 amino acids) and three contained free carboxyl-terminal amino acids. These toxins (plectoxins) are paralytic and/or lethal when injected into insect pests such as the lepidopterous larvae H. virescens, Spodoptera exigua (beet armyworm), and Manduca sexta (tobacco hornworm). The authors expect these plectoxins to be useful in insect control after delivery by a suitable agent such as a recombinant baculovirus.

IT 151066-05-6 151066-06-7 151066-09-0 155807-95-7
RL: BIOL (Biological study)
(amino acid sequence and paralytic and/or letal for insects of)

IT 155807-94-6
RL: BIOL (Biological study)
(partial amino acid sequence and paralytic and/or letal for insects of)

L13 ANSWER 21 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1993:643503 CAPLUS
DOCUMENT NUMBER: 119:243503
TITLE: Isolation of insecticidal toxins from Plectreurys tristis
INVENTOR(S): Leisy, Douglas J.; Quistad, Gary B.; Skinner, Wayne S.
PATENT ASSIGNEE(S): Sandoz AG, Switz.; Sandoz-Patent-G.m.b.H.; Sandoz-Erfindungen Verwaltungsgesellschaft m.b.H.
SOURCE: Eur. Pat. Appl., 50 pp.
CODEN: EPXXDW
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE

EP 556160	A2	19930818	EP 1993-810078	19930208
EP 556160	A3	19931027		
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LI, NL, PT, SE				
JP 05310791	A2	19931122	JP 1993-22725	19930210
US 5470735	A	19951128	US 1994-221285	19940330
US 6265376	B1	20010724	US 1995-428596	19950425
PRIORITY APPLN. INFO.:			US 1992-837194 A	19920211
			US 1993-58051 B1	19930503
			US 1993-163602 B1	19931206
			US 1994-221285 A3	19940330

ED Entered STN: 11 Dec 1993

AB Novel plectoxins (Plt) are isolated from the venom of primitive hunting spider, Plectreurys tristis, and their amino acid sequences disclosed. The amino acid sequences of the plectoxins bear remarkable homol., especially with respect to the positioning of the Cys residue. The amino acid sequence of the most potent toxin Plt-VI is used for cloning its cDNA by PCR. Baculovirus-based expression vectors for Plt-VI are constructed and their insecticidal activities are demonstrated.

IT 151066-05-6P, Plectoxin PltVI (Plectreurys tristis)
151066-06-7P, Plectoxin PltVIII (Plectreurys tristis)
151066-09-0P, Plectoxin PltXI (Plectreurys tristis)
RL: PREP (Preparation)
(amino acid sequence of and purification of)

L13 ANSWER 22 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1993:575022 CAPLUS
DOCUMENT NUMBER: 119:175022
TITLE: Nucleotide sequence of the genes for ribosomal proteins HS15 and HSH from Haloarcula marismortui: an archaeon-specific gene cluster
AUTHOR(S): Arndt, Evelyn; Steffens, Christina
CORPORATE SOURCE: Abt. Wittmann, Max-Planck-Inst. Mol. Genet., Berlin, D-1000/33, Germany
SOURCE: FEBS Letters (1992), 314(3), 211-14
CODEN: FEBLAL; ISSN: 0014-5793
DOCUMENT TYPE: Journal
LANGUAGE: English

ED Entered STN: 30 Oct 1993

AB The nucleotide sequences of the genes for two ribosomal proteins, HS15 and HSH, from the archaeon H. marismortui, have been determined The genes were found in a cluster together with another open reading frame with a probable regulatory function. HS15 and HSH have counterparts in eukaryotes. HS15 is significantly homologous to S19 from frog (Xenopus laevis). HSH is related to S37 from yeast (Saccharomyces cerevisiae) and S27 from fly (Drosophila melanogaster), as well as to other members of the S27 family. Eubacterial counterparts were not found, suggesting that these proteins are extra proteins that are absent in eubacterial ribosomes.

IT 149348-68-5
RL: BAC (Biological activity or effector, except adverse); BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)
(amino acid sequence of, complete)

L13 ANSWER 23 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 1993:403384 CAPLUS
DOCUMENT NUMBER: 119:3384
TITLE: Metal binding properties of single amino acid deletion mutants of zinc finger peptides: Studies using cobalt(II) as a spectroscopic probe
AUTHOR(S): Shi, Yigong; Beger, Richard D.; Berg, Jeremy M.

CORPORATE SOURCE: Sch. Med., Johns Hopkins Univ., Baltimore, MD, 21205, USA
SOURCE: Biophysical Journal (1993), 64(3), 749-53
CODEN: BIOJAU; ISSN: 0006-3495
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 10 Jul 1993
AB Peptides correspond to Cys2His2 zinc finger domains from which one amino acid has been deleted have been synthesized and their metal-binding properties characterized. In contrast to earlier reports (Parraga, G., et al., 1990), such peptides do bind metal ions such as cobalt(II). A peptide with the sequence ProTyrLysCysProGluCysLysSerPheSerGlnLysSerAspLeu ValLysHisGlnArgThrHisThrGly (which corresponds to a previously characterized consensus zinc finger sequence from which a Gly residue immediately following the second Cys residue has been deleted) was found to form a 1:1 peptide to cobalt(II) complex with an absorption spectrum quite similar to those previously observed for zinc finger peptide-cobalt(II) complexes. The dissociation constant for this complex is 6 + 10-6 M, a factor of 100 times higher than that for the parent peptide. A peptide with the sequence LysProTyrProCysGlyLeuCysArgCysPheThrArgArgAspLeuLeuIleAr gHisAlaGlnLysIleHisSerGlyAsnLeu corresponding to a similar mutation of the peptide ADR1 was also characterized. Spectroscopic studies with cobalt(II) revealed that this peptide forms both 1:1 and 2:1 peptide to cobalt(II) complexes. The absorption spectra of the two forms and the dissociation consts. were determined via deconvolution methods. In contrast, the parent peptide ADR1a was found to form only a 1:1 complex under comparable conditions and this 1:1 complex was found to be more stable than that for the mutant. These results reveal that deletion mutations do adversely affect the stability of zinc finger peptide-metal complexes but that the effects are not as drastic as had been previously described.
IT [illegible]
RL: BIOL (Biological study)
(zinc binding by, cobalt in study of, zinc finger structure of proteins in relation to)

L13 ANSWER 24 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1992:627297 CAPLUS
DOCUMENT NUMBER: 117:227297
TITLE: Complete nucleotide sequence of the virus SSV1 of the archaebacterium Sulfolobus shibatae
AUTHOR(S): Palm, Peter; Schleper, Christa; Grampp, Bernd; Yeats, Siobhan; McWilliam, Peter; Reiter, Wolf Dieter; Zillig, Wolfram
CORPORATE SOURCE: Max-Planck-Inst. Biochem., Martinsried, 8033, Germany
SOURCE: Virology (1991), 185(1), 242-50
CODEN: VIRLAX; ISSN: 0042-6822
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 13 Dec 1992
AB The DNA sequence of the S. shibatae virus SSV1 is the first complete sequence of an archaebacterial virus genome. The viral DNA is a closed double-stranded DNA circle of 15465 bp. The features of the sequences, the positions of all 11 transcripts, the 3 characterized proteins, and the open reading frames are described.
IT [illegible]Peptide (bacteriophage SSV1 5.55-kilodalton reduced)
RL: PRP (Properties)
(amino acid sequence of)

L13 ANSWER 25 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1992:566108 CAPLUS
DOCUMENT NUMBER: 117:166108

TITLE: Aromatic-aromatic interactions in the zinc finger motif. Analysis of the two-dimensional nuclear magnetic resonance structure of a mutant domain
AUTHOR(S): Jasanoff, Alan; Kochoyan, Michel; Fraenkel, Ernest; Lee, Jonathan P.; Weiss, Michael A.
CORPORATE SOURCE: Dep. Biol. Chem. Mol. Pharmacol., Harvard Med. Sch., Boston, MA, 02115, USA
SOURCE: Journal of Molecular Biology (1992), 225(4), 1035-47
CODEN: JMOBAK; ISSN: 0022-2836
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 01 Nov 1992
AB The folding and stability of globular proteins are determined by a variety of chemical mechanisms, including hydrogen bonds, salt bridges and the hydrophobic effect. Of particular interest are weakly polar interactions involving aromatic rings, which are proposed to regulate the geometry of closely packed protein interiors. Such interactions reflect the electrostatic contribution of π-electrons and, unlike van der Waals' interactions and the hydrophobic effect, may, in principle, introduce a directional force in a protein's hydrophobic core. Although the weakly polar hypothesis is supported by a statistical anal. of protein structures, the general importance of such contributions to protein folding and stability is unclear. Here, the presence of alternative aromatic-aromatic interactions in the two-dimensional NMR structure of a mutant Zn finger is shown. Changes in aromatic packing lead in turn to local and non-local differences between the structures of a wild-type and mutant domain. The results provide insight into the evolution of Zn finger sequences and have implications for understanding how geometric relationships may be chemical encoded in a simple sequence template.
IT 136444-42-3 137924-60-8 143863-57-4
RL: PRP (Properties)
(conformation of aromatic amino acid residues interaction in, NMR study of, in protein zinc finger motif modeling)

L13 ANSWER 26 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1992:506628 CAPLUS
DOCUMENT NUMBER: 117:106628
TITLE: Two-dimensional NMR studies of the zinc finger motif: solution structure and dynamics of mutant ZFY domains containing aromatic substitutions in the hydrophobic core
AUTHOR(S): Qian, Xiuqi; Weiss, Michael A.
CORPORATE SOURCE: Dep. Biol. Chem. Mol. Pharmacol., Harvard Med. Sch., Boston, MA, 02115, USA
SOURCE: Biochemistry (1992), 31(33), 7463-76
CODEN: BICHAW; ISSN: 0006-2960
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 20 Sep 1992
AB Solution structures of mutant Zn fingers containing aromatic substitutions in the hydrophobic core are determined by 2D-NMR spectroscopy and distance-geometry/simulated annealing (DG/SA). The wild-type domain (designated ZFY-6) is derived from the human male-associated protein ZFY and represents a sequence motif (Cys-X2-Cys-X-Ar-X7-Leu-X2-His-X4-His) that differs in the location (aromatic swap) and diversity (Ar = tyrosine, phenylalanine, or histidine) of the central aromatic residue from the consensus sequence. In a given ZFY domain the choice of a particular aromatic residue is invariant among vertebrates, suggesting that alternative swapped aromatic residues are functionally inequivalent. 2D-NMR studies of analogs containing tyrosine, phenylalanine, or histidine at the swapped site were conducted. The 3 DG/SA structures each retain the ββα motif and exhibit

similar staggered-horizontal packing between the variant aromatic residue and the proximal histidine in the hydrophobic core. The structures and stabilities of the tyrosine and phenylalanine analogs are essentially identical, differing only by local exposure of polar (Tyr p-OH) or nonpolar (Phe p-H) surfaces. The dynamic stability of the histidine analog is reduced as indicated by more rapid proton-deuterium exchange of hydrogen bonds related to secondary structure and amide-sulfur coordination (slowly exchanging amide resonances in D2O) and by more extensive averaging of main-chain dihedral angles (3JαNH coupling consts.). An aspartic acid in the putative DNA recognition surface, whose configuration is well-defined as a possible helix N-cap in the tyrosine and phenylalanine analogs, exhibits multiple weak main-chain contacts in the NOESY spectrum of the histidine analog; such NOEs are geometrically inconsistent and so provide complementary evidence for structural fluctuations. Because the 3 DG ensembles have similar apparent precision, the finding of reduced dynamic stability in the histidine analog emphasizes the importance of expts. that directly probe fluctuations at several time scales. The results provide insight into the design of biol. metal-binding sites and suggest that the dynamics of a protein surface may contribute to DNA recognition.

IT 136444-42-3P 142439-45-0P 142439-46-1P
RL: SPN (Synthetic preparation); PREP (Preparation)
(preparation and conformation and dynamics of)

L13 ANSWER 27 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1992:17349 CAPLUS
DOCUMENT NUMBER: 116:17349
TITLE: Architectural rules of the zinc-finger motif: comparative two-dimensional NMR studies of native and "aromatic-swap" domains define a "weakly polar switch"
AUTHOR(S): Kochoyan, Michel; Keutmann, Henry T.; Weiss, Michael A.
CORPORATE SOURCE: Dep. Biol. Chem. Mol. Pharmacol., Harvard Med. Sch., Boston, MA, 02115, USA
SOURCE: Proceedings of the National Academy of Sciences of the United States of America (1991), 88(19), 8455-9
CODEN: PNASA6; ISSN: 0027-8424
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 24 Jan 1992
AB The Zn-finger motif, encoding a globular minidomain with characteristic structure, provides a striking example of a sequence template for protein folding. Insight into architectural rules relating the amino acid sequence of a protein to its structure and stability may be obtained by comparative study of analogs. As the first step toward defining such rules for the Zn finger, the design of an aromatic-swap analog based in the ZFY two-finger repeat (a conserved alternation in sequence pattern observed among odd- and even-numbered domains in a family of sex-related vertebrate transcription factors) was recently described. Consensus and swapped aromatic residues, introduced as revertants of less stable aromaticless analogs, provided equivalent contributions to the thermodn. stability of the Zn finger. Here the solution structures of a wild-type domain and an aromatic-swap analog, as determined by two-dimensional NMR and distance-geometry/restrained mol. dynamics calcns. are described and compared. The wild-type and aromatic-swap analog each contain an N-terminal β-sheet and a C-terminal α-helix (ββα motif),
as observed in other systems, and exhibit a highly ordered hydrophobic core in which the native or swapped aromatic ring is closely packed. Remarkably, however, the two structures are stabilized by alternative aromatic-aromatic interactions, which in turn alter the resp. DNA-binding surfaces. The results suggest that native and swapped Zn-finger sequences encode a

weakly polar switch between thermodynamically equivalent but functionally distinct architectures for DNA recognition.

IT [redacted]
RL: PRP (Properties)
(conformation of, internal packing and surface architecture in, aromatic-aromatic interactions in, DNA-binding zinc-finger proteins in relation to)

L13 ANSWER 28 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1991:577687 CAPLUS
DOCUMENT NUMBER: 115:177687
TITLE: Alternating zinc fingers in the human male-associated protein ZFY: HX3H and HX4H motifs encode a local structural switch
AUTHOR(S): Kochoyan, Michel; Keutmann, Henry T.; Weiss, Michael A.
CORPORATE SOURCE: Dep. Biol. Chem. Mol. Pharmacol., Harvard Med. Sch., Boston, MA, 02115, USA
SOURCE: Biochemistry (1991), 30(39), 9396-402
CODEN: BICHAW; ISSN: 0006-2960
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 01 Nov 1991
AB The two-finger repeat in the human male-associated protein ZFY provides a model for comparative 2-dimensional (2D) NMR studies of classical and variant Zn fingers. This repeat is defined in part by an alternation in spacing between consensus (HX3H) and variant (HX4H) histidine spacings. To investigate the effects of a switch between alternative histidine spacings, an HX3H analog of a representative HX4H domain of known structure [ZFY-6] was designed. The HX3H analog (designated ZFY-switch) forms a tetrahedral Co2+ complex whose thermodn. stability is similar to that of the parent peptide. The 2D-NMR studies demonstrate that ZFY-switch and ZFY-6, although similar in overall structure, exhibit significant local changes near the site of deletion. Whereas the HX4H site in the native finger forms a nonstandard loop, the HX3H site in the ZPY-switch folds as a 310 extension of the C-terminal α-helix, as observed in the NMR solution structure of a consensus HX3H domain and the crystal structure of a representative Zn finger-DNA complex. It is proposed that variant histidine spacings (HX3H and HX4H) encode a local switch between alternative surface architectures with implications for models of protein-DNA recognition.
IT [redacted]
RL: PRP (Properties)
(conformation of, structural switch in alternating zinc fingers of human male associated protein ZFY in relation to)

L13 ANSWER 29 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1991:444501 CAPLUS
DOCUMENT NUMBER: 115:44501
TITLE: Alternating zinc fingers in the human male associated protein ZFY: refinement of the NMR structure of an even finger by selective deuterium labeling and implications for DNA recognition
AUTHOR(S): Kochoyan, Michel; Keutmann, Henry T.; Weiss, Michael A.
CORPORATE SOURCE: Dep. Biol. Chem. Mol. Pharmacol., Harvard Med. Sch., Boston, MA, 02115, USA
SOURCE: Biochemistry (1991), 30(29), 7063-72
CODEN: BICHAW; ISSN: 0006-2960
DOCUMENT TYPE: Journal
LANGUAGE: English

ED Entered STN: 10 Aug 1991
AB ZFY, a male-associated Zn-finger protein encoded by the human Y chromosome, exhibits a distinctive 2-finger repeat: whereas odd-numbered domains fit a general consensus, even-numbered domains exhibit systematic differences. Do these odd and even sequences encode structurally distinct surfaces for DNA recognition. As a 1st step toward answering this question, the authors have recently described the sequential 1H NMR assignment of a representative nonconsensus Zn finger (designated ZFY-6T) based on 2D NMR studies of a 30-residue peptide (Kochoyan, M., et al., 1991). Initial structural modeling by distance geometry/stimulated annealing (DG/SA) demonstrated that this peptide retained the N-terminal β-hairpin and C-terminal α-helix (ββα motif) observed in consensus Zn fingers. However, the precision of this initial structure was limited by resonance overlap, which led to ambiguities in the assignment of key NOEs in the hydrophobic core. In this paper these ambiguities are resolved by selective deuterium labeling, enabling a refined structure to be calculated by DG/SA and restrained mol. dynamics. These calcns. provide a detailed view of the hydrophobic core and protein surface, which are analyzed in reference to previously characterized Zn fingers. Variant (even) and consensus (odd) aromatic residues Y10 and F12, shown in an aromatic swap analog to provide equivalent contributions to the hydrophobic core, nevertheless exhibit striking differences in packing interactions: Y10-but not F12-contributes to a contiguous region of the protein surface defined by putative specificity-determining residues. Alternating surface architectures may have implications for the mechanism of DNA recognition by the ZFY 2-finger repeat.
IT 134629-32-6
RL: BIOL (Biological study)
(conformation and other structural characteristics of, gene ZFY protein of human structure in relation to)

L13 ANSWER 30 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1990:586790 CAPLUS
DOCUMENT NUMBER: 113:186790
TITLE: Alternating zinc finger motifs in the male-associated protein ZFY: defining architectural rules by mutagenesis and design of an "Aromatic Swap" second-site revertant
AUTHOR(S): Weiss, Michael A.; Keutmann, Henry T.
CORPORATE SOURCE: Dep. Biol. Chem. Mol. Pharmacol., Harvard Med. Sch., Boston, MA, 02115, USA
SOURCE: Biochemistry (1990), 29(42), 9808-13
CODEN: BICHAW; ISSN: 0006-2960
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 23 Nov 1990
AB Spectroscopic and biochem. studies of native and mutant Zn finger peptides from ZFY, a putative transcription factor encoded by the sex-determining region of the human Y chromosome are described. The parent peptide, based on ZFY domain 6, exhibits metal-dependent helix formation within a rigid tertiary framework. Surprisingly, nonarom. substitution of the consensus aromatic group (Tyr10→Ser or Lys) is compatible with native architecture but results in loss of stability to pH or guanidine denaturation. Remarkably, these perturbations are reverted by a second-site mutation in which an alternative aromatic residue is introduced (Ser12→Phe). Design of the second-site revertant (aromatic swap) is based on the ZFY 2-finger repeat, a conserved symmetry among the ZFY-related Zn finger proteins, and is in accord with recent 2-dimensional NMR structures of Zn finger peptides. These expts. suggest general rules for metal-dependent folding of the Zn finger motif.
IT 127570-37-0 127570-37-0D, cobalt and zinc complexes

129731-17-5 129731-17-5D, cobalt complexes
129756-45-2 129756-45-2D, cobalt complexes
129785-68-8 129785-68-8D, cobalt complexes
129785-69-9 129785-69-9D, cobalt complexes
RL: BIOL (Biological study)
(conformation and stability of, ZFY protein in relation to)

L13 ANSWER 31 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1990:547476 CAPLUS
DOCUMENT NUMBER: 113:147476
TITLE: High-resolution three-dimensional structure of a single zinc finger from a human enhancer binding protein in solution
AUTHOR(S): Omichinski, James G.; Clore, G. Marius; Appella, Ettore; Sakaguchi, Kazuyasu; Gronenborn, Angela M.
CORPORATE SOURCE: Lab. Chem. Phys., Natl. Inst. Diabetes Dig. Kidney Dis., Bethesda, MD, 20892, USA
SOURCE: Biochemistry (1990), 29(40), 9324-34
CODEN: BICHAW; ISSN: 0006-2960
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 27 Oct 1990
AB The 3-dimensional structure of a 30-residue synthetic peptide containing the C-terminal zinc finger motif of a human enhancer-binding protein [called human immunodeficiency virus type I enhancer-binding protein (HIV-EPI), major histocompatibility complex-binding protein (MBP-1), or pos. regulatory domain II of the human interferon β promoter-binding factor 1 (PRDII-BF1)] was determined by 2-dimensional NMR spectroscopy and hybrid distance geometry-dynamical simulated annealing calcns. The structure determination is based on 487 approx. interproton distance and 63 torsion angle (.vphi., ψ, and χ) restraints. A total of 40 simulated annealing structures were calculated, and the atomic rms distribution about the mean coordinate positions (excluding residues 29 and 30, which are ill-defined) is 0.4 Å for the backbone atoms, 0.8 Å for all atoms, and 0.41 Å for all atoms excluding the lysine and arginine side chains, which are disordered. The solution structure of the zinc finger consists of 2 irregular antiparallel β-strands connected by an atypical turn (residues 3-12) and a classical α-helix (residues 14-24). The zinc is tetrahedrally coordinated to the S atoms of 2 cysteines (Cys-5 and Cys-8) and to the Nε2 atoms of 2 histidines (His-21 and His-27). The 2 cysteine residues are located in the turn connecting the 2 β-strands (residues 5-8); one of the histidine ligands (His-21) is in the α-helix, whereas the second histidine (His-27) is at the end of a looplike structure (formed by the end of the α-helix and a turn). The general architecture is qual. similar to 2 previously determined low-resolution Cys2His2 zinc finger structures, although distinct differences can be observed in the β-strands and turn and in the region around the 2 histidines coordinated to zinc. Comparison of the overall polypeptide fold of the enhancer-binding protein zinc finger with known structures in the crystallog. data base reveals a striking similarity to one region (residues 23-44) of the x-ray structure of proteinase inhibitor domain III of Japanese quail ovomucoid (Papamokos, E., et al., 1982), which could be superimposed with a backbone atomic rms difference of 0.95 Å on residues 3-25 (excluding residue 6) of the zinc finger from the enhancer-binding protein. The presence of structural homol. between 2 proteins of very different function may indicate that the so-called zinc finger motif is not unique for a class of DNA binding proteins but may represent a general folding motif found in a variety of proteins irresp. of their function.
IT 129193-43-7D, complexes with zinc
RL: PRP (Properties)

(structure of, as model for zinc finger domain of human enhancer-binding protein)

L13 ANSWER 32 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1990:419708 CAPLUS
DOCUMENT NUMBER: 113:19708
TITLE: Alternating zinc-finger motifs in the human male-associated protein ZFY
AUTHOR(S): Weiss, Michael A.; Mason, Kathleen A.; Dahl, Charles E.; Keutmann, Henry T.
CORPORATE SOURCE: Dep. Biol. Chem. Mol. Pharmacol., Harvard Med. Sch., Boston, MA, 02115, USA
SOURCE: Biochemistry (1990), 29(24), 5660-4
CODEN: BICHAW; ISSN: 0006-2960
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 21 Jul 1990
AB ZFY, a putative transcription factor encoded by the human Y chromosome, contains a distinctive 2-finger repeat: odd-numbered and even-numbered CC/HH metal-binding motifs exhibit systematic alternation in sequence pattern. Such alternation, which is not generally observed in Zn-finger proteins, has also been described in an extensive family of Kruppel-like genes in Xenopus laevis and in the AIDS-associated human DNA-binding protein HIV-EP1. The strict conservation of a 2-finger repeat among ZFY-, Kruppel-, and HIV-related Zn-finger proteins suggests distinct mechanisms of protein-nucleic acid recognition. To test whether this sequence pattern reflects an underlying alternation in domain structure, single-finger peptides from the human ZFY gene were synthesized and characterized. Remarkably, systematic differences in metal-dependent folding are observed in the CD spectra of even- and odd-numbered domains. The results suggest the existence of distinct CC/HH finger submotifs, which may play different roles in nucleic acid recognition.
IT 127570-37-0P 127570-38-1P 127570-39-2P
RL: RCT (Reactant); SPN (Synthetic preparation); PREP (Preparation); RACT (Reactant or reagent)
(preparation and conformation of, as models of human protein ZFY)

L13 ANSWER 33 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1990:231428 CAPLUS
DOCUMENT NUMBER: 112:231428
TITLE: The arrangement of disulfide loops in human α2-HS glycoprotein. Similarity to the disulfide bridge structures of cystatins and kininogens
AUTHOR(S): Kellermann, Josef; Haupt, Heinz; Auerswald, Ernst August; Mueller-Esterl, Werner
CORPORATE SOURCE: Max-Planck-Inst. Biochem., Martinsried, D-8033, Fed. Rep. Ger.
SOURCE: Journal of Biological Chemistry (1989), 264(24), 14121-8
CODEN: JBCHA3; ISSN: 0021-9258
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 23 Jun 1990
AB The complete SS loop structure of human α2-HS glycoprotein has been elucidated. α2-HS glycoprotein isolated from human plasma was found to be a 2-chain protein composed of a heavy and a light chain. The heavy chain comprises the A-chain of α2-HS glycoprotein and part of the connecting peptide which has been predicted from the corresponding cDNA sequence, whereas the light chain corresponds to the B-chain of α2-HS glycoprotein. Twelve half-cystine residues are present in the α2-HS glycoprotein mol., and 11 of them are positioned in the heavy

chain and a single one in the light chain of the mol.; they form 6 SS bridges. The 1st and the last half-cystine residues of the amino acid sequence of α2-HS glycoprotein are engaged in the formation of a loop spanning the extreme N- and C-terminal portions of the mol., thereby connecting the heavy and light chains. The other 10 half-cystines residues are linked consecutively in the heavy chain and form 5 loops which span 4-19 amino acid residues. Among them are 2 pairs of loops which are characterized by mutual sequence homol. The particular arrangement of SS loops in α2-HS glycoprotein is similar to the patterns of linearly arranged and tandemly repeated SS loops of cysteine proteinase inhibitors, i.e., the cystatins and the kininogens. It is concluded that α2-HS glycoprotein represents a structural prototype of a novel family among the cystatin superfamily, characterized by the presence of 2 cystatin-like building blocks. Extensive similarity among the N-terminal sequences of α2-HS glycoprotein and human histidine-rich glycoprotein suggest that the latter protein is another candidate protein of this new family.

IT 12727-92-1, Glucoprotein α2HS (human protein moiety)
RL: PRP (Properties)
(amino acid sequence of)

L13 ANSWER 34 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1990:153886 CAPLUS
DOCUMENT NUMBER: 112:153886
TITLE: Spectroscopic studies of wild-type and mutant "zinc finger" peptides: determinants of domain folding and structure
AUTHOR(S): Parraga, Grace; Horvath, Suzanne; Hood, Leroy; Young, E. T.; Klevit, Rachel E.
CORPORATE SOURCE: Dep. Biochem., Univ. Washington, Seattle, WA, 98195, USA
SOURCE: Proceedings of the National Academy of Sciences of the United States of America (1990), 87(1), 137-41
CODEN: PNASA6; ISSN: 0027-8424
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 28 Apr 1990
AB The Zn finger model (Miller, J., et al., 1985; Brown, R. S., et al., 1985) makes both specific structural and specific functional predictions about Zn-finger consensus sequences that can be tested with a combination of genetic, mol. biol., and biophys. techniques. The yeast transcription factor ADR1 contains 2 adjacent Zn finger domains; genetic and deletion analyses showed that amino acid substitutions and deletions in the Zn finger domains resulted in the loss of protein activity. To test the structural and folding predictions of the Zn finger model, peptides encompassing each of the ADR1 fingers were synthesized (ADR1a and ADR1b) as well as a mutant finger peptide (del138) deleted for a single amino acid residue. The folding and metal-binding characteristics of these were assessed by 1H NMR and visible spectroscopy. While a single unique conformational species was detected for the 2 wild-type peptides upon tetrahedral binding of Zn, the deletion peptide did not bind Zn with tetrahedral geometry, nor did it fold into a Zn finger domain. The metal-binding and folding results found with the mutant peptide were similar to those obtained when thiol alkylation or imidazole protonation of the wild-type peptides was performed. These data indicate that ligand spacing and both thiol and imidazole participation in Zn binding are specific and necessary requirements for Zn finger folding, which provides direct support for the initial predictions of the model.

IT 126182-06-7
RL: PRP (Properties)
(zinc binding by and zinc-finger conformation of, structural

determinants of)

L13 ANSWER 35 OF 35 CAPLUS COPYRIGHT 2004 ACS on STN
ACCESSION NUMBER: 1990:17879 CAPLUS
DOCUMENT NUMBER: 112:17879
TITLE: The position of the disulfide bonds in human plasma α2HS-glycoprotein and the repeating double disulfide bonds in the domain structure
AUTHOR(S): Araki, Tomohiro; Yoshioka, Yasuyuki; Schmid, Karl
CORPORATE SOURCE: Sch. Med., Boston Univ., Boston, MA, USA
SOURCE: Biochimica et Biophysica Acta (1989), 994(3), 195-9
CODEN: BBACAQ; ISSN: 0006-3002
DOCUMENT TYPE: Journal
LANGUAGE: English
ED Entered STN: 21 Jan 1990
AB The positions of the inter- and intrachain SS bonds of human plasma α2HS-glycoprotein were determined α2HS-glycoprotein was digested with acid proteinase and then with thermolysin. SS bond-containing peptides were separated by reversed-phase HPLC and detected by the SBD-F (7-fluorobenzo-2-oxa-1,3-diasole-4-sulfonic acid ammonium salt) method. One inter-SS bond-containing peptide and 5 intra-SS bond-containing peptides (A-chain) were purified and the cysteines involved were identified as Cys-18 (B-chain)-Cys-14 (A-chain), Cys-71-Cys-82, Cys-71-Cys-82, Cys96-Cys-114, Cys-128-Cys-131, Cys-190-Cys-201, and Cys-212-Cys-229, resp. The location of the intra-SS bonds revealed that the A-chain of α2HS-glycoprotein is composed of 3 domains. Two domains possess intramol. homol. judging from the total chain length of the domains, size of the loops formed by the SS bonds, the location of 2 SS loops near the C-terminal end of domains A and B, the distance between 2 SS bonds of each domain, the amino acid sequence homol. between these 2 domains (22.6%), number of amino acid residues between the 2nd SS loops and the end of domains A and B, and the positions of the ordered structures.
IT **124146-20-9P**, Glycoprotein α2HS (human blood plasma protein moiety)
RL: SPN (Synthetic preparation); PREP (Preparation)
(preparation of)

=> sel hit rn l13 1-35 ⏎ # Hit Registry #'s selected out of references, crossed back into Registry, & combined with seq search answer set
E1 THROUGH E44 ASSIGNED

=> => d his l14

(FILE 'REGISTRY' ENTERED AT 13:07:06 ON 13 DEC 2004)

FILE 'CAPLUS' ENTERED AT 13:07:06 ON 13 DEC 2004
SEL HIT RN L13 1-35

FILE 'REGISTRY' ENTERED AT 13:08:34 ON 13 DEC 2004
L14 44 S E1-E44 AND L3 ⏎

=> d cn kwic nte l14 1-44 ⏎ fil hom

L14 ANSWER 1 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN GenBank AAV47294 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAV47294 (Translated from: GenBank AY596297)
RN **775190-06-2** REGISTRY
SQL **44** ← use Registry # to match sequence to citation
SQL = Sequence length
SEQ 1 MPHNEYYNDD GELDRETCPR CGDTVLAEHE DRQHCGKCGY TEWK

====== ========== ========== ==
HITS AT: 15-42

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L14 ANSWER 2 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN GenBank AAV45792 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAV45792 (Translated from: GenBank AY596297)
RN [redacted] REGISTRY
SQL **44**

SEQ 1 MVVQTERDEV MWYKCETCGL MFDDQNDARQ HEENCDDEDP SYIQ
========= ========== =========
HITS AT: 12-39

L14 ANSWER 3 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN GenBank AAV44349 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAV44349 (Translated from: GenBank AY596292)
RN [redacted] REGISTRY
SQL **44**

SEQ 1 MAIEKRGNAV IATGCPFCDS DIDPQQSLAK HLNNDCPEFG GESA
========= ========== =========
HITS AT: 12-39

L14 ANSWER 4 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Threonine, L-alanyl-L-alanylglycyl-L-valyl-L-seryl-L-lysyl-L-asparaginyl-L-arginyl-L-cysteinyl-L-lysyl-L-phenylalanyl-L-cysteinyl-L-α-aspartyl-L-lysyl-L-asparaginyl-L-phenylalanyl-L-alanyl-L-leucyl-L-α-glutamyl-L-arginyl-L-alanyl-L-leucyl-L-histidyl-L-isoleucyl-L-histidyl-L-leucyl-L-methionyl-L-glutaminyl-L-asparaginyl-L-cysteinyl-L-α-aspartyl-L-lysyl-L-isoleucyl-L-prolyl-L-prolyl-L-alanyl-L-α-glutamyl-L-lysyl-L-arginyl-L-lysyl-L-leucyl-L-α-glutamyl-L-phenylalanyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 48: PN: US6703491 SEQID: 60048 claimed protein
CN Protein (Drosophila melanogaster clone US6703491-SEQID-60048 fragment)
RN [redacted] REGISTRY
SQL **44**

SEQ 1 AAGVSKNRCK FCDKNFALER ALHIHLMQNC DKIPPAEKRK LEFT
===== ========== ========== ===
HITS AT: 6-33

L14 ANSWER 5 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Tryptophan, L-valyl-L-glutaminyl-L-isoleucyl-L-histidyl-L-arginyl-L-cysteinyl-L-tyrosyl-L-serylglycyl-L-tyrosyl-L-arginyl-L-cysteinyl-L-leucyl-L-arginyl-L-cysteinyl-L-cysteinyl-L-glutaminyl-L-alanyl-L-tryptophyl-L-threonyl-L-prolylglycyl-L-cysteinyl-L-seryl-L-prolyl-L-cysteinyl-L-leucyl-L-histidyl-L-cysteinyl-L-phenylalanyl-L-phenylalanyl-L-tryptophyl-L-phenylalanyl-L-threonyl-L-leucyl-L-leucyl-L-leucyl-L-tryptophyl-L-leucyl-L-threonyl-L-prolyl-L-leucyl-L-leucyl-L-leucyl-L-prolyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 2428: PN: US6703491 SEQID: 58428 claimed protein
CN Protein (Drosophila melanogaster clone US6703491-SEQID-58428 fragment)
RN [redacted] REGISTRY
SQL **46**

SEQ 1 VQIHRCYSGY RCLRCCQAWT PGCSPCLHCF FWFTLLLWLT PLLLPW
== ========== ========== ======
HITS AT: 9-36

L14 ANSWER 6 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Valine, L-arginyl-L-asparaginyl-L-phenylalanyl-L-cysteinyl-L-tryptophyl-L-tryptophyl-L-cysteinyl-L-arginyl-L-asparaginyl-L-arginyl-L-seryl-L-threonyl-L-threonyl-L-prolylglycyl-L-leucylglycyl-L-alanyl-L-threonyl-L-histidyl-L-prolyl-L-prolyl-L-threonyl-L-alanyl-L-leucyl-L-arginyl-L-leucyl-L-leucyl-L-leucyl-L-seryl-L-serylglycyl-L-tyrosyl-L-threonyl-L-asparaginyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 1180: PN: US6703491 SEQID: 41180 claimed protein
CN Protein (Drosophila melanogaster clone US6703491-SEQID-41180 fragment)
RN 669255-25-8 REGISTRY
SQL 36

SEQ 1 RNFCWWCRNR STTPGLGATH PPTALRLLLS SGYTNV
========== ========== ========
HITS AT: 1-28

L14 ANSWER 7 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Phenylalanine, L-seryl-L-alanyl-L-leucyl-L-lysyl-L-arginyl-L-histidyl-L-leucyl-L-α-glutamyl-L-phenylalanyl-L-α-glutamyl-L-cysteinylglycyl-L-methionyl-L-leucyl-L-α-glutamyl-L-asparaginyl-L-phenylalanyl-L-arginyl-L-cysteinyl-L-glutaminyl-L-valyl-L-cysteinyl-L-α-aspartyl-L-alanylglycyl-L-phenylalanyl-L-lysyl-L-arginyl-L-lysyl-L-α-aspartyl-L-seryl-L-leucyl-L-asparaginyl-L-arginyl-L-histidyl-L-cysteinyl-L-lysyl-L-valyl-L-lysyl-L-lysyl-L-histidyl-L-asparaginyl-L-threonyl-L-lysyl-L-tyrosyl-L-leucyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 654: PN: US6703491 SEQID: 40654 claimed protein
CN Protein (Drosophila melanogaster clone US6703491-SEQID-40654 fragment)
RN 669250-26-4 REGISTRY
SQL 47

SEQ 1 SALKRHLEFE CGMLENFRCQ VCDAGFKRKD SLNRHCKVKK HNTKYLF
===== ========== ========== ===
HITS AT: 16-43

L14 ANSWER 8 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Isoleucine, L-leucyl-L-lysyl-L-phenylalanyl-L-cysteinylglycyl-L-isoleucyl-L-cysteinyl-L-cysteinyl-L-α-aspartyl-L-lysyl-L-cysteinyl-L-asparaginyl-L-cysteinyl-L-valyl-L-prolyl-L-seryl-L-arginyl-L-threonyl-L-tyrosyl-L-arginyl-L-isoleucyl-L-isoleucyl-L-α-aspartyl-L-histidyl-L-seryl-L-prolyl-L-cysteinyl-L-tyrosyl-L-threonyl-L-α-aspartyl-L-methionyl-L-isoleucyl-L-asparaginyl-L-threonyl-L-asparaginylglycyl-L-asparaginyl-L-α-aspartyl-L-lysyl-L-cysteinyl-L-threonyl-L-tryptophyl-L-threonyl-L-arginyl-L-tyrosyl-L-lysyl-L-phenylalanyl-L-isoleucyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 2929: PN: US20040031072 SEQID: 178929 claimed protein
CN Transcription-associated protein (Glycine max clone PAT_MRT3847_13258C.1.pep fragment)
RN 666902-20-1 REGISTRY
SQL 49

SEQ 1 LKFCGICCDK CNCVPSRTYR IIDHSPCYTD MINTNGNDKC TWTRYKFII
====== ========== ========== ==
HITS AT: 5-32

L14 ANSWER 9 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Tyrosine, L-α-aspartyl-L-valyl-L-phenylalanyl-L-arginyl-L-tyrosyl-L-phenylalanyl-L-alanyl-L-leucyl-L-seryl-L-tyrosyl-L-prolyl-L-cysteinyl-L-valyl-L-tryptophyl-L-tryptophyl-L-cysteinyl-L-lysyl-L-prolyl-L-cysteinyl-L-cysteinyl-L-histidyl-L-arginyl-L-histidyl-L-tyrosyl-L-methionyl-L-α-glutamyl-L-α-aspartyl-L-valyl-L-phenylalanyl-L-alanyl-L-seryl-L-histidyl-L-α-aspartyl-L-alanyl-L-threonyl-L-methionyl-L-leucyl-L-α-glutamyl-L-lysyl-L-valyl-L-valyl-L-leucyl-L-cysteinyl-L-α-glutamyl-L-α-glutamyl-L-leucyl-L-phenylalanyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 3840: PN: US20040031072 SEQID: 143840 claimed protein
CN Transcription-associated protein (Glycine max clone PAT_MRT3847_100900C.1.pep fragment)
RN 665228-20-6 REGISTRY
SQL 48

SEQ 1 DVFRYFALSY PCVWWCKPCC HRHYMEDVFA SHDATMLEKV VLCEELFY
======== ========== ==========
HITS AT: 13-40

L14 ANSWER 10 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN GenBank AAH46856 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAH46856 (TRANSLATED FROM: GenBank BC046856)
RN 623621-65-8 REGISTRY
SQL 34

SEQ 1 MGATCHACIG GTNVQNEMLG HLTLLGRQDE CSTC
========= ========== =========
HITS AT: 2-29

L14 ANSWER 11 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN Protein (Sulfolobus spindle-shaped virus strain K1 open reading frame ORF C43) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAQ94382
CN GenBank AAQ94382 (Translated from: GenBank AY423772)
RN 606872-49-5 REGISTRY
SQL 43

SEQ 1 MYQCLRCGGI FRKRREVVEH LLSGHMQSKF TLEYFYVYFR VRE
========== ========== ========
HITS AT: 1-28

L14 ANSWER 12 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN GenBank AAQ73267 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAQ73267 (Translated from: GenBank AY370762)
CN Protein (Sulfolobus spindle-shaped virus 2 ORF 48)
RN 583804-75-5 REGISTRY
SQL 48

SEQ 1 MMLMYQCLRC GSIFDKRSEV IEHLLSVHGQ MNKVTLEYFY IYFKVRRP
======= ========== ========== =
HITS AT: 4-31

L14 ANSWER 13 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-phenylalanyl-L-alanyl-L-cysteinyl-L-prolyl-L-α-glutamyl-L-cysteinyl-L-prolyl-L-lysyl-L-arginyl-L-phenylalanyl-L-methionyl-L-glutaminyl-L-arginyl-L-seryl-L-alanyl-L-leucyl-L-threonyl-L-valyl-L-

histidyl-L-threonyl-L-threonyl-L-lysyl-L-leucyl-L-histidyl-L-prolyl-L-asparaginyl-L-lysyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 29

SEQ 1 KFACPECPKR FMQRSALTVH TTKLHPNKK
========== ========== ========
HITS AT: 1-28

L14 ANSWER 14 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Histidine, L-methionyl-L-alanyl-L-alanyl-L-leucyl-L-asparaginyl-L-leucyl-L-α-glutamyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-alanyl-L-histidyl-L-cysteinyl-L-valyl-L-valyl-L-arginyl-L-phenylalanyl-L-tyrosyl-L-valyl-L-asparaginyl-L-α-glutamyl-L-asparaginyl-L-lysyl-L-leucyl-L-tryptophyl-L-histidyl-L-isoleucyl-L-α-aspartyl-L-seryl-L-methionyl-L-α-aspartyl-L-arginyl-L-tyrosyl-L-lysyl-L-arginyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAF82870
CN GenBank AAF82870 (Translated from: GenBank AE003859)
CN Protein (Xylella fastidiosa gene XF0057)
RN [redacted] REGISTRY
SQL 36

SEQ 1 MAALNLEYQC AHCVVRFYVN ENKLWHIDSM DRYKRH
==== ========== ========== ====
HITS AT: 7-34

L14 ANSWER 15 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN 41: PN: WO0033068 FIGURE: 29 unclaimed protein (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 46

SEQ 1 LCGRRLILFM LATCGECDTX XXDSSEDLSH ICCIKGCDVQ DIIRVC
========== ========== ========
HITS AT: 11-38
NTE

type	location		description
uncommon	Aaa-20	-	-
uncommon	Aaa-21	-	-
uncommon	Aaa-22	-	-

L14 ANSWER 16 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN 131: PN: WO0033068 SEQID: 131 unclaimed protein (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 46

SEQ 1 LCGRRLILFM LATCGECDTX XXDSSEDLSH ICCIKQCDVQ DIIRVC
========== ========== ========
HITS AT: 11-38
NTE

type	location		description
uncommon	Aaa-20	-	-
uncommon	Aaa-21	-	-
uncommon	Aaa-22	-	-

L14 ANSWER 17 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Arginine, L-prolyl-L-tyrosyl-L-lysyl-L-cysteinyl-L-glutaminyl-L-leucyl-L-cysteinyl-L-tyrosyl-L-tyrosyl-L-α-glutamyl-L-threonyl-L-lysyl-L-histidyl-L-threonyl-L-α-glutamyl-L-α-glutamyl-L-leucyl-L-α-aspartyl-L-seryl-L-histidyl-L-leucyl-L-arginyl-L-asparaginyl-L-α-glutamyl-L-histidyl-L-lysyl-L-valyl-L-seryl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL **29**

SEQ 1 PYKCQLCYYE TKHTEELDSH LRNEHKVSR
========== ========== ========
HITS AT: 1-28

L14 ANSWER 18 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN Glycine, L-methionyl-L-lysyl-L-valyl-L-histidyl-L-methionyl-L-histidyl-L-threonyl-L-lysyl-L-phenylalanyl-L-cysteinyl-L-leucyl-L-isoleucyl-L-cysteinyl-L-leucyl-L-leucyl-L-threonyl-L-phenylalanyl-L-isoleucyl-L-phenylalanyl-L-histidyl-L-histidyl-L-cysteinyl-L-asparaginyl-L-histidyl-L-cysteinyl-L-histidyl-L-α-glutamyl-L-α-glutamyl-L-histidyl-L-α-aspartyl-L-histidylglycyl-L-prolyl-L-α-glutamyl-L-alanyl-L-leucyl-L-histidyl-L-arginyl-L-glutaminyl-L-glutaminyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN Secretory protein (human clone 33-10-4-G2-PU precursor N-terminal fragment)
RN [redacted] REGISTRY
SQL **41**

SEQ 1 MKVHMHTKFC LICLLTFIFH HCNHCHEEHD HGPEALHRQQ G
==== ========== ========== ====
HITS AT: 7-34

L14 ANSWER 19 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Proline, L-glutaminyl-L-tyrosyl-L-arginyl-L-cysteinyl-L-alanyl-L-α-aspartyl-L-cysteinyl-L-α-aspartyl-L-tyrosyl-L-alanyl-L-threonyl-L-lysyl-L-tyrosyl-L-cysteinyl-L-histidyl-L-seryl-L-phenylalanyl-L-lysyl-L-leucyl-L-histidyl-L-leucyl-L-arginyl-L-histidyl-L-tyrosylglycyl-L-histidyl-L-lysyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL **28**

SEQ 1 QYRCADCDYA TKYCHSFKLH LRHYGHKP
========== ========== ========
HITS AT: 1-28

L14 ANSWER 20 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Serine, L-alanyl-L-valyl-L-lysyl-L-cysteinyl-L-isoleucylglycyl-L-tryptophyl-L-glutaminyl-L-α-glutamyl-L-threonyl-L-cysteinyl-L-asparaginylglycyl-L-glutaminyl-L-leucyl-L-prolyl-L-cysteinyl-L-cysteinyl-L-α-aspartylglycyl-L-cysteinyl-L-valyl-L-methionyl-L-cysteinyl-L-α-glutamyl-L-cysteinyl-L-asparaginyl-L-isoleucyl-L-methionylglycyl-L-glutaminyl-L-asparaginyl-L-cysteinyl-L-arginyl-L-cysteinyl-L-asparaginyl-L-histidyl-L-prolyl-L-lysyl-L-methionyl-L-threonyl-L-seryl-L-α-glutamyl-L-cysteinylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN Plectoxin Plt-VIII (Plectreurys tristes clone pSCI273 reduced)
CN Toxin Plt VIII (Plectreurys tristes clone pSCI273 reduced)
RN [redacted] REGISTRY
SQL **46**

SEQ 1 AVKCIGWQET CNGQLPCCDG CVMCECNIMG QNCRCNHPKM TSECGS
=== ========== ========== =====
HITS AT: 18-45

L14 ANSWER 21 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Serine, L-alanyl-L-valyl-L-lysyl-L-cysteinyl-L-isoleucylglycyl-L-tryptophyl-L-glutaminyl-L-α-glutamyl-L-threonyl-L-cysteinyl-L-asparaginylglycyl-L-lysyl-L-leucyl-L-prolyl-L-cysteinyl-L-cysteinyl-L-α-aspartylglycyl-L-cysteinyl-L-valyl-L-methionyl-L-cysteinyl-L-α-glutamyl-L-cysteinyl-L-asparaginyl-L-isoleucyl-L-methionylglycyl-L-glutaminyl-L-asparaginyl-L-cysteinyl-L-arginyl-L-cysteinyl-L-asparaginyl-L-histidyl-L-prolyl-L-lysyl-L-alanyl-L-threonyl-L-seryl-L-α-glutamyl-L-cysteinyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN 155807-95-7 REGISTRY
SQL 46

SEQ 1 AVKCIGWQET CNGKLPCCDG CVMCECNIMG QNCRCNHPKA TSECES
=== ========== ========== =====
HITS AT: 18-45

L14 ANSWER 22 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Glutamic acid, L-alanyl-L-valyl-L-lysyl-L-cysteinyl-L-isoleucylglycyl-L-tryptophyl-L-glutaminyl-L-α-glutamyl-L-threonyl-L-cysteinyl-L-asparaginylglycyl-L-asparaginyl-L-leucyl-L-prolyl-L-cysteinyl-L-cysteinyl-L-asparaginyl-L-α-glutamyl-L-cysteinyl-L-valyl-L-methionyl-L-cysteinyl-L-α-glutamyl-L-cysteinyl-L-asparaginyl-L-isoleucyl-L-methionylglycyl-L-glutaminyl-L-asparaginyl-L-cysteinyl-L-arginyl-L-cysteinyl-L-asparaginyl-L-histidyl-L-prolyl-L-lysyl-L-alanyl-L-threonyl-L-asparaginyl-L-α-glutamyl-L-cysteinyl- (9CI) (CA INDEX NAME)
RN 155807-94-6 REGISTRY
SQL 45

SEQ 1 AVKCIGWQET CNGNLPCCNE CVMCECNIMG QNCRCNHPKA TNECE
=== ========== ========== =====
HITS AT: 18-45

L14 ANSWER 23 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Serine, L-α-glutamyl-L-valyl-L-lysyl-L-cysteinyl-L-isoleucylglycyl-L-tryptophyl-L-glutaminyl-L-α-glutamyl-L-tyrosyl-L-cysteinyl-L-arginylglycyl-L-asparaginyl-L-leucyl-L-prolyl-L-cysteinyl-L-cysteinyl-L-α-aspartyl-L-α-aspartyl-L-cysteinyl-L-valyl-L-methionyl-L-cysteinyl-L-α-glutamyl-L-cysteinyl-L-asparaginyl-L-isoleucyl-L-methionylglycyl-L-glutaminyl-L-asparaginyl-L-cysteinyl-L-arginyl-L-cysteinyl-L-asparaginyl-L-histidyl-L-prolyl-L-arginyl-L-isoleucyl-L-threonyl-L-seryl-L-α-glutamyl-L-cysteinylglycyl- (9CI) (CA INDEX NAME)
OTHER NAMES:
CN Plectoxin Plt-XI (Plectreurys tristes clone pSCI265 reduced)
CN Plectoxin PltXI (Plectreurys tristis)
CN Toxin Plt XI (Plectreurys tristes clone pSCI265 reduced)
RN 151066-09-0 REGISTRY
SQL 46

SEQ 1 EVKCIGWQEY CRGNLPCCDD CVMCECNIMG QNCRCNHPRI TSECGS
=== ========== ========== =====
HITS AT: 18-45

L14 ANSWER 24 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Serine, L-alanyl-L-valyl-L-lysyl-L-cysteinyl-L-isoleucylglycyl-L-tryptophyl-L-glutaminyl-L-α-glutamyl-L-threonyl-L-cysteinyl-L-asparaginylglycyl-L-lysyl-L-leucyl-L-prolyl-L-cysteinyl-L-cysteinyl-L-

α-aspartylglycyl-L-cysteinyl-L-valyl-L-methionyl-L-cysteinyl-L-α-glutamyl-L-cysteinyl-L-asparaginyl-L-isoleucyl-L-methionylglycyl-L-glutaminyl-L-asparaginyl-L-cysteinyl-L-arginyl-L-cysteinyl-L-asparaginyl-L-histidyl-L-prolyl-L-lysyl-L-methionyl-L-threonyl-L-seryl-L-α-glutamyl-L-cysteinylglycyl- (9CI) (CA INDEX NAME)

OTHER NAMES:
CN Plectoxin PltVIII (Plectreurys tristis)
RN 151066-06-7 REGISTRY
SQL **46**

SEQ 1 AVKCIGWQET CNGKLPCCDG CVMCECNIMG QNCRCNHPKM TSECGS
=== ========== ========== =====
HITS AT: 18-45

L14 ANSWER 25 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Serine, L-alanyl-L-valyl-L-lysyl-L-cysteinyl-L-isoleucylglycyl-L-tryptophyl-L-glutaminyl-L-α-glutamyl-L-threonyl-L-cysteinyl-L-asparaginylglycyl-L-asparaginyl-L-leucyl-L-prolyl-L-cysteinyl-L-cysteinyl-L-asparaginyl-L-α-glutamyl-L-cysteinyl-L-valyl-L-methionyl-L-cysteinyl-L-α-glutamyl-L-cysteinyl-L-asparaginyl-L-isoleucyl-L-methionylglycyl-L-glutaminyl-L-asparaginyl-L-cysteinyl-L-arginyl-L-cysteinyl-L-asparaginyl-L-histidyl-L-prolyl-L-lysyl-L-alanyl-L-threonyl-L-asparaginyl-L-α-glutamyl-L-cysteinyl-L-α-glutamyl- (9CI) (CA INDEX NAME)

OTHER NAMES:
CN Plectoxin Plt-VI (Plectreurys tristes clones pSCI263 and pSCI266 reduced)
CN Plectoxin Plt-VI (Plectreurys tristes clones pSCI267,pSCI268,pSCI269,pSCI270,pSCI271,pSCI272 reduced)
CN Plectoxin PltVI (Plectreurys tristis)
CN Toxin Plt VI (Plectreurys tristes clone pSCI263 reduced)
CN Toxin Plt VI (Plectreurys tristes clone pSCI266 reduced)
RN 151066-05-6 REGISTRY
SQL **46**

SEQ 1 AVKCIGWQET CNGNLPCCNE CVMCECNIMG QNCRCNHPKA TNECES
=== ========== ========== =====
HITS AT: 18-45

L14 ANSWER 26 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-methionyl-L-prolyl-L-histidyl-L-asparaginyl-L-α-glutamyl-L-tyrosyl-L-tyrosyl-L-asparaginyl-L-α-aspartyl-L-α-aspartylglycyl-L-α-glutamyl-L-leucyl-L-α-aspartyl-L-arginyl-L-α-glutamyl-L-threonyl-L-cysteinyl-L-prolyl-L-arginyl-L-cysteinylglycyl-L-α-aspartyl-L-threonyl-L-valyl-L-leucyl-L-alanyl-L-α-glutamyl-L-histidyl-L-α-glutamyl-L-α-aspartyl-L-arginyl-L-glutaminyl-L-histidyl-L-cysteinylglycyl-L-lysyl-L-cysteinylglycyl-L-tyrosyl-L-threonyl-L-α-glutamyl-L-tryptophyl- (9CI) (CA INDEX NAME)

OTHER NAMES:
CN Protein HSH (Haloarcula marismortui ribosome precursor reduced)
RN 149348-68-5 REGISTRY
SQL **44**

SEQ 1 MPHNEYYNDD GELDRETCPR CGDTVLAEHE DRQHCGKCGY TEWK
====== ========== ========== ==
HITS AT: 15-42

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L14 ANSWER 27 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN Peptide (bacteriophage SSV1 5.55-kilodalton reduced) (9CI) (CA INDEX

NAME)
RN 144350-77-6 REGISTRY
SQL 45

SEQ 1 MYQCLRCGGI FNKRREVVEH LLVGHKHKDR LTLDFYYIYF RVRGQ
========== ========== ========

HITS AT: 1-28

RELATED SEQUENCES AVAILABLE WITH SEQLINK

L14 ANSWER 28 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-threonyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-tyrosyl-L-arginyl-L-phenylalanyl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN 143863-57-4 REGISTRY
SQL 30

SEQ 1 KTYQCQYCEY RFADSSNLKT HIKTKHSKEK
========== ========== ==========

HITS AT: 2-29

L14 ANSWER 29 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-threonyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-histidyl-L-arginyl-L-seryl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN 142439-46-1 REGISTRY
SQL 30

SEQ 1 KTYQCQYCEH RSADSSNLKT HIKTKHSKEK
========== ========== ==========

HITS AT: 2-29

L14 ANSWER 30 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-threonyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-phenylalanyl-L-arginyl-L-seryl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN 142439-45-0 REGISTRY
SQL 30

SEQ 1 KTYQCQYCEF RSADSSNLKT HIKTKHSKEK
========== ========== ==========

HITS AT: 2-29

L14 ANSWER 31 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-threonyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-lysyl-L-arginyl-L-phenylalanyl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN 137924-60-8 REGISTRY
SQL 30

SEQ 1 KTYQCQYCEK RFADSSNLKT HIKTKHSKEK
========= ========== =========
HITS AT: 2-29

L14 ANSWER 32 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-threonyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-tyrosyl-L-arginyl-L-seryl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 30

SEQ 1 KTYQCQYCEY RSADSSNLKT HIKTKHSKEK
========= ========== =========
HITS AT: 2-29

L14 ANSWER 33 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN Zincate(3-), [L-lysyl-L-threonyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-tyrosyl-L-arginyl-L-seryl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysylglycyl-L-lysinato(5-)]-, trihydrogen, (T-4)- (9CI) (CA INDEX NAME)
OTHER CA INDEX NAMES:
CN L-Lysine, L-lysyl-L-threonyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-tyrosyl-L-arginyl-L-seryl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysylglycyl-, zinc complex
RN [redacted] REGISTRY
SQL 30

SEQ 1 KTYQCQYCEY RSADSSNLKT HIKTKHSKGK
========= ========== =========
HITS AT: 2-29

RELATED SEQUENCES AVAILABLE WITH SEQLINK
NTE metal complex

L14 ANSWER 34 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-prolyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-tyrosyl-L-arginyl-L-phenylalanyl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 30

SEQ 1 KPYQCQYCEY RFADSSNLKT HIKTKHSKEK
========= ========== =========
HITS AT: 2-29

L14 ANSWER 35 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-prolyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-lysyl-L-arginyl-L-phenylalanyl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY

SQL 30

SEQ 1 KPYQCQYCEK RFADSSNLKT HIKTKHSKEK
========== ========== ==========

HITS AT: 2-29

L14 ANSWER 36 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-prolyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-lysyl-L-arginyl-L-seryl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 30

SEQ 1 KPYQCQYCEK RSADSSNLKT HIKTKHSKEK
========== ========== ==========

HITS AT: 2-29

L14 ANSWER 37 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-prolyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-seryl-L-arginyl-L-seryl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 30

SEQ 1 KPYQCQYCES RSADSSNLKT HIKTKHSKEK
========== ========== ==========

HITS AT: 2-29

L14 ANSWER 38 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-arginyl-L-prolyl-L-tyrosyl-L-histidyl-L-cysteinyl-L-seryl-L-tyrosyl-L-cysteinyl-L-asparaginyl-L-phenylalanyl-L-seryl-L-phenylalanyl-L-lysyl-L-threonyl-L-lysylglycyl-L-asparaginyl-L-leucyl-L-threonyl-L-lysyl-L-histidyl-L-methionyl-L-lysyl-L-seryl-L-lysyl-L-alanyl-L-histidyl-L-seryl-L-lysyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 30

SEQ 1 RPYHCSYCNF SFKTKGNLTK HMKSKAHSKK
========== ========== ==========

HITS AT: 2-29

L14 ANSWER 39 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-valyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-tyrosyl-L-seryl-L-threonyl-L-threonyl-L-α-aspartyl-L-alanyl-L-serylglycyl-L-phenylalanyl-L-lysyl-L-arginyl-L-histidyl-L-valyl-L-isoleucyl-L-seryl-L-isoleucyl-L-histidyl-L-threonyl-L-lysyl-L-α-aspartyl- (9CI) (CA INDEX NAME)
RN [redacted] REGISTRY
SQL 30

SEQ 1 KVYQCQYCEY STTDASGFKR HVISIHTKDK
========== ========== ==========

HITS AT: 2-29

L14 ANSWER 40 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-threonyl-L-histidyl-L-glutaminyl-L-cysteinyl-L-leucyl-L-histidyl-L-cysteinyl-L-α-aspartyl-L-histidyl-L-lysyl-L-seryl-L-

seryl-L-asparaginyl-L-seryl-L-seryl-L-α-aspartyl-L-leucyl-L-lysyl-L-arginyl-L-histidyl-L-valyl-L-isoleucyl-L-seryl-L-valyl-L-histidyl-L-threonyl-L-lysyl-L-α-aspartyl- (9CI) (CA INDEX NAME)
RN [illegible] REGISTRY
SQL **30**

SEQ 1 KTHQCLHCDH KSSNSSDLKR HVISVHTKDK
========= ========== =========
HITS AT: 2-29

L14 ANSWER 41 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Lysine, L-lysyl-L-prolyl-L-tyrosyl-L-glutaminyl-L-cysteinyl-L-glutaminyl-L-tyrosyl-L-cysteinyl-L-α-glutamyl-L-tyrosyl-L-arginyl-L-seryl-L-alanyl-L-α-aspartyl-L-seryl-L-seryl-L-asparaginyl-L-leucyl-L-lysyl-L-threonyl-L-histidyl-L-isoleucyl-L-lysyl-L-threonyl-L-lysyl-L-histidyl-L-seryl-L-lysyl-L-α-glutamyl- (9CI) (CA INDEX NAME)
RN [illegible] REGISTRY
SQL **30**

SEQ 1 KPYQCQYCEY RSADSSNLKT HIKTKHSKEK
========= ========== =========
HITS AT: 2-29

L14 ANSWER 42 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN Glycoprotein α2HS (human protein moiety) (9CI) (CA INDEX NAME)
RN [illegible] REGISTRY
SQL 309,282,**27**

SEQ 101 LKLDGKFSVV YAKCDSSPDS AEDVRKVCQD CPLLAPLNDT RVVHAAKAAL
====== ========== ==========
151 AAFNAQNNGS NFQLEEISRA QLVPLPPSTY VEFTVSGTDC VAKEATEAAK
==
HITS AT: 125-152

NTE multichain

type	------ location ------		description
bridge	Cys-14	- Cys-18'	disulfide bridge
bridge	Cys-71	- Cys-82	disulfide bridge
bridge	Cys-96	- Cys-114	disulfide bridge
bridge	Cys-128	- Cys-131	disulfide bridge
bridge	Cys-190	- Cys-201	disulfide bridge
bridge	Cys-212	- Cys-229	disulfide bridge

L14 ANSWER 43 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN L-Leucine, L-lysyl-L-prolyl-L-tyrosyl-L-prolyl-L-cysteinylglycyl-L-leucyl-L-cysteinyl-L-asparaginyl-L-arginyl-L-cysteinyl-L-phenylalanyl-L-threonyl-L-arginyl-L-arginyl-L-α-aspartyl-L-leucyl-L-leucyl-L-isoleucyl-L-arginyl-L-histidyl-L-alanyl-L-glutaminyl-L-lysyl-L-isoleucyl-L-histidyl-L-serylglycyl-L-asparaginyl- (9CI) (CA INDEX NAME)
RN [illegible] REGISTRY
SQL **30**

SEQ 1 KPYPCGLCNR CFTRRDLLIR HAQKIHSGNL
========= ========== =========
HITS AT: 2-29

L14 ANSWER 44 OF 44 REGISTRY COPYRIGHT 2004 ACS on STN
CN Glycoprotein α2HS (human blood plasma protein moiety) (9CI) (CA

INDEX NAME)
RN [redacted] REGISTRY
SQL 309,282,27

SEQ 101 LKLDGKFSVV YAKCDSSPDS AEDVRKVCQD CPLLAPLNDT RVVHAAKAAL
====== ========== ==========
151 AAFNAQNNGS NFQLEEISRA QLVPLPPSTY VEFTVSGTDC VAKEATEAAK
==
HITS AT: 125-152

NTE multichain

type	------ location ------		description
bridge	Cys-14	- Cys-18'	disulfide bridge
bridge	Cys-71	- Cys-82	disulfide bridge
bridge	Cys-96	- Cys-114	disulfide bridge
bridge	Cys-128	- Cys-131	disulfide bridge
bridge	Cys-190	- Cys-201	disulfide bridge
bridge	Cys-212	- Cys-229	disulfide bridge

FILE 'HOME' ENTERED AT 13:09:04 ON 13 DEC 2004

=>